NATIONAL CONVENTION TRANSACTIONS 1957

ELEVENTH ANNUAL CONVENTION

AMERICAN SOCIETY FOR QUALITY CONTROL

NATIONAL CONVENTION TRANSACTIONS 1957

ELEVENTH ANNUAL CONVENTION

AMERICAN SOCIETY FOR QUALITY CONTROL

MAY 22, 23, 24, 1957

DETROIT, MIGHIGAN

PRICE

$3.00 at the Convention

$3.50 if Ordered by Mail

Copyright 1957, By
AMERICAN SOCIETY FOR QUALITY CONTROL, INC.
Room 6197
Plankinton Building
161 West Wisconsin Avenue
Milwaukee 3, Wisconsin

Printed in the United States of America By
UNIVERSAL LITHOGRAPHERS, INC.
Baltimore, Md.

STRATEGY IN RESEARCH

D. S. McArthur and J. J. Heigl
Esso Research and Engineering Company

Abstract

This paper discusses strategy in empirical research. Several research teams have done empirical research on electrical analogs of research problems where a random, normally distributed error was introduced to measurements of the dependent variable. The success attained with different strategies has been observed. Designed experiments are "better" than the one-variable-at-a-time strategy. However, the "best" strategy is not known.

The work pointed up mistakes we are apt to make in doing research:

We aren't always bold enough;
We are apt to get in a rut;
We are too easily fooled by the error;
We don't always eliminate the unimportant variables first; and
We don't always know when to quit.

Some of the "elements" which will probably be found in the "best" research strategy are suggested.

I. DOING "RESEARCH ON RESEARCH"

This paper is concerned with strategy in empirical research. We have done some "research on research" using simple electrical analogs of research problems. The effectiveness of different empirical strategies, used by different research teams, has been compared. So far, the research has simply consisted of observing how the different strategies seemed to work out. As the work progresses, we hope to delineate research strategies in sufficient detail to program them on a high-speed computer. In this way, a particular research strategy could be tried out a large number of times (on different kinds of problems) until an "expected payoff" for that strategy could be obtained. Work of this type (with the help of the mathematicians to suggest different strategies) may develop new methods of doing empirical research which are more economical and effective than our present methods.

II. THE "LITTLE BLACK BOXES"

A research problem can be thought of as a "little black box". The box has dials (variables) on it which affect a meter reading (the dependent variable). The problem is to increase (let's say) the meter reading. If we are to do basic research on the problem, we take the box apart and study each component in detail. We learn exactly how each part works and how it interacts with the other parts to produce a change in the meter reading. If we are to do empirical research, we study the box from the outside. We move the dials and observe what happens to the meter.

We have made "little black boxes" and done empirical research on them. In these boxes, five dials control a meter not seen by the researcher. He runs experiments on the system and gets test results just as he would on a real problem. In real problems the results contain an "error". This is because only some of the many factors affecting the system are being controlled. The uncontrolled effects make the dependent measurement subject to random fluctuations (error). The random fluctuations were introduced, in these models, by the plant "operator". He superimposed random variations on the meter reading in accordance with values drawn at random from a deck of cards containing 1000 normally distributed deviations.

The analogs used in these studies are simple internally. They are electrical resistance networks powered by batteries. Dials indicated the settings on variable resistors. The output was a voltmeter reading.

The units were housed in two electrically connected boxes: One contained the variable dials and was used by the experimenters, the other contained the error dial and the voltmeter and was used by the plant operator. The boxes were oriented so that the research workers could not see the voltmeter (i.e., they were not permitted to optimize by twiddling the dials with an eye on the voltmeter). Figure 1 shows a three-man team doing research on one of these systems called the ALCOHOL PLANT.

DOING RESEARCH ON THE ALCOHOL PLANT

FIGURE 1

A. The "ALCOHOL PLANT"

This model represented a hypothetical alcohol plant having a yield of 24%. The yield was affected by five variables A, B, C, D, and E. Two versions of the ALCOHOL PLANT were used in experimental work. They were identical outwardly but differed in internal wiring. The first model had relatively simple relationships between the variables (not many interactions) but a large unknown error (standard deviation of 4%, absolute value) in the yield measurement. The true relationships between the variables were as shown in Figure 2. There was an interaction between variable B and E. Variable D had no effect on the system.

The second version had more complex relationships between the variables as shown in Figure 3, but had a smaller known error (standard deviation of 2.5%, absolute value) in the yield measurement.

In both ALCOHOL PLANT models it was desired to increase the yield from the plant by doing empirical research work on it.

B. The SNAFU Unit

A small amount of work was done on another model called the SNAFU Unit containing six variables. This was intended to simulate a pilot plant unit which would be available for study for a limited period of time only. It was desired to get the maximum amount of information on how to operate the unit in this fixed time (enough time for 40 test runs). Four of the six variables were continuous (like those in the ALCOHOL PLANTS), two of them were discreet variables, one corresponding to three different feed stocks and the other to three different catalysts. There was considerable interaction between all variables. The error in yield response was unknown and relatively small (standard deviation of 1.5%, absolute value).

C. How The Games Were Played

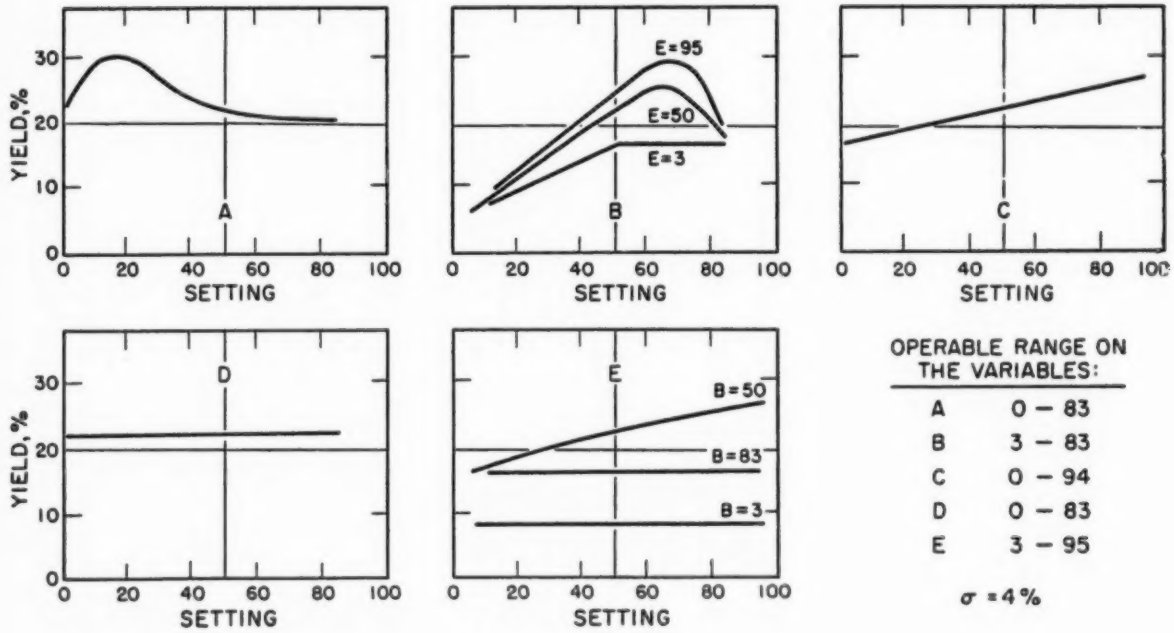
It is apparent that these models represented problems which are much simpler than real problems. The operating space was clearly defined -- it simply consisted of the range on each dial (resistor). There were no islands or inlets in operating space. The operating range on a variable was independent of the settings on the other variables. There were only a relatively few variables to consider in these problems. Most real problems would have more. There were no second-order interactions introduced into the models (although this was not known to the research workers). The economics were clean-cut; much more so than would be the case in real problems.

However, the problems were sufficiently difficult to give many research teams trouble. We have been able to learn, even from these simplified problems, some of the mistakes we make in doing research and what type of approach seems the best.

In the ALCOHOL PLANT models we provided a set of mythical economics. Each plant test cost $20,000. A 1% improvement in yield is worth $500,000 ($50,000 per year over the 10-year life of the plant). Each variable had a known operating range from about 0 to 85 dial reading. At the start of the research work each variable was set at 50 and this

FIG. 2

FIRST ALCOHOL PLANT
— EFFECT OF THE VARIABLES
(TAKEN ONE AT A TIME —OTHERS SET AT 50)



OPERABLE RANGE ON THE VARIABLES:

A	0 – 83
B	3 – 83
C	0 – 94
D	0 – 83
E	3 – 95

σ = 4 %

gave a known yield of 24%. The research workers were told that the maximum possible yield from the plant was something less than 50% but they weren't told what it was (it might be 25% or 49%).

Research work was done on these models by individuals or by research teams competing against each other. The research team agreed on a strategy and decided on the plant tests which they wished to run. The plant tests were run off with the team making the dial settings and the operator introducing the random error. The plant result (yield) was read off to the research team by the plant operator. Usually where more than one played as a team, they were given time to work out their strategy. Many of the games were played by teams who had a week in which to work out their strategy and make their calculations. In the case of many individual players, there was little time to work out a strategy. The player spent less than an hour on the research work.

III. HOW SHOULD WE COMPARE RESEARCH STRATEGIES?

In the work done on the ALCOHOL PLANTS, we were faced with the problem of deciding on a criterion to use in comparing research strategies. Three criteria suggest themselves.

A. Best Yield From a Fixed Budget?

We might give each research team a fixed budget, say enough for 40 tests, and measure their performance on the basis of the yield improvement they got with this budget. This means of measuring success was used in the case of the SNAFU Unit. It has the drawback that we do not know how big to make this fixed budget. Forty tests may be altogether too few, considering the large return possible on the investment. On the other hand, we may be encouraging waste. Possibly most of the gain can be obtained in 10 or 20 tests and further refinements may not be worthwhile. This criterion does not leave the research worker with the difficult decision of when to quit research.

B. Highest Return on The Dollar?

We could compare strategies on the basis of the payoff per dollar spent. The strategy which produced (on the average) the highest return per dollar spent would then be judged the best one. If the researcher got a yield improvement of 10%, for example, after running 50 tests, his payoff (using the economics for this plant) would be five million dollars for a research cost of one million dollars so that the return on research is five dollars per dollar spent.

This has the disadvantage that by some stroke of luck (or good management) a team might get a yield improvement of 2% after running only four tests. The payoff would then be $12.50 per dollar spent. If this team knew enough to quit research at this point, they might come away with the prize, even though there is still a lot for the Company to gain by upping the plant yield further. This does not seem like the best way for management to judge research efficiency.

C. Biggest Payoff Over Competitive Research?

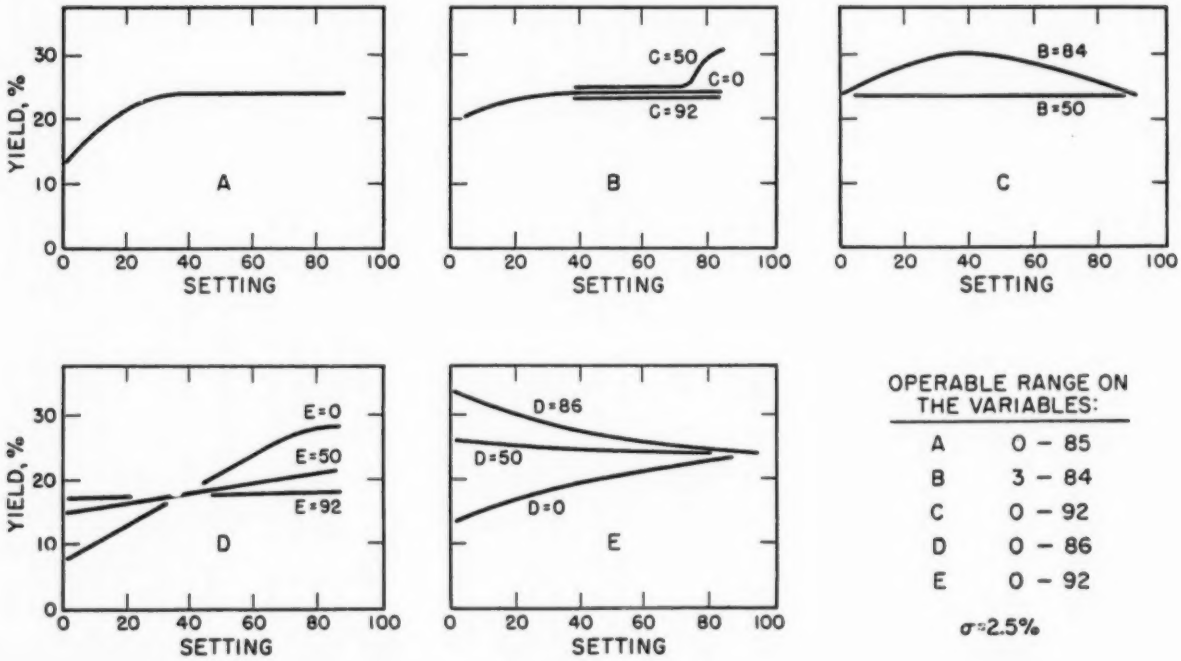
In the ALCOHOL PLANT games, we decided that the best research

FIG. 3

SECOND ALCOHOL PLANT

— EFFECT OF THE VARIABLES

(TAKEN ONE AT A TIME — OTHERS SET AT 50)



OPERABLE RANGE ON THE VARIABLES:	
A	0 – 85
B	3 – 84
C	0 – 92
D	0 – 86
E	0 – 92

σ=2.5%

strategy is the one showing the greatest return above that which could be obtained in competitive research. Competitive research represents the average of all other research work. It has been said that research on the average will net about $5 dollars per dollar spent, e.g., a return of $6 per dollar spent. The best strategy is then the one which will give the biggest payoff beyond this competitive research level.

On this basis, the team obtaining a yield improvement of 10% in 50 tests (which made a gross profit of five million dollars but spent one million dollars on research) had a net payoff of only four million dollars, whereas, competitive research (on the average) would have netted five million. This team lost money! They made one million dollars less than could have been made by spending the same money in competitive research.

The team making a 2% improvement in yield after running four tests made money, but not as much as they might have. They netted $920,000, whereas, competitive research on the average would have netted only $400,000 so that they beat competitive research to the tune of $520,000. However, further research might have increased this margin over competitive research. The best research team will continue research until the payoff drops to the competitive payoff level.

IV. COMPARISON OF RESEARCH STRATEGIES USED

Sixty-one research teams experimented with these simulated research problems. Some of these teams consisted of one man only, others consisted of two or three men competing as a team. The competitors came from widely different fields of interest. Many of them were industrial research workers, many were administrators (primarily from research), some were mathematicians, and some were new college graduates. Many different strategies were used.

A. One-Variable-At-A-Time

By far the most common strategy employed was that of changing one variable at a time while holding the others constant. Thirty-four out of the 61 teams used essentially the one-variable-at-a-time technique. This strategy was employed by almost all administrators, by new college graduates and by most of the research workers who had not had training in the design of experiments. The one-variable-at-a-time strategy seems to be the accepted way of doing research.

B. Using Designed Experiments

Many of the research workers had been trained in experimental design, although only a few had had practical experience in this field. The prevalance of the designed experiment approach is undoubtedly due to this training. Twenty-one out of the 61 research teams made use of some sort of designed experiment. By far the most common type of experimental design was the fractional factorial. Most people used a $1/4$ replicate of a 2^5 factorial experiment for their first design. Lack of experience in the design of experiments and analysis of the data led to some difficulty in interpretation of the data.

C. Box-Wilson Technique

In at least two cases a serious effort was made to employ the Box-Wilson technique. This was done without previous experience in the use of this technique so that the results may not be truly representative of the payoff which should be expected in using this strategy.

D. Eliminating Operating Space

Two or three teams employed the concept of elimination of operating space. The first experiment in space was chosen through the use of random numbers. When this was found to be an unsatisfactory place in operating space, random numbers were again used to choose a point in the remaining operating space. After each experiment, a volume was eliminated from the operating space and random numbers were used to choose a point in the remaining operable space. This was continued until a good place was found.

E. Other Techniques

A few other techniques were tried, such as changing two variables at a time and employing the concept of randomizing interaction vectors between the variables.

F. Comparison of the Strategies

The results obtained by the various research teams on these models are listed in Tables I, II, and III. These briefly describe the strategy used and show the result obtained in terms of the profit over competitive research. The results on the first ALCOHOL PLANT model show that people using the one-test-at-a-time strategy did as well, on the average, as those using some type of designed experiment. Luck seemed to play a major role in the success of the research teams. Most of the researchers had no idea of the size of the error. Very few of them discovered that variable D had no effect on the system. There was a general feeling of helplessness and frustration. Many of the teams felt that the error (standard deviation of 4%, absolute value) was ridiculously large, when they heard what it was. Actually, it is relatively no larger than the errors in many real problems on which the research teams had worked. It is the feeling of the authors that researchers experienced in the design of experiments would average better than those using the one-variable-at-a-time strategy on this model in the long run. Experiments run by the authors using designed experiments generally showed a high return.

On the second model (which had a smaller, known error) people using some form of designed experiment did much better, on the average, than those using the one-test-at-a-time strategy. The 16 teams using the one-variable-at-a-time strategy lost 1.1 million dollars, on the average, whereas the 12 teams using the designed experiment approach made 0.8 million dollars on the average. Only 2 of the 16 teams using the one-variable-at-a-time strategy made money; 6 of the 12 teams using the designed experiment approach made money. Luck still played a part in the success attained but it seemed to be less important here than in the first ALCOHOL PLANT model.

It was harder to show a real gain in the second ALCOHOL PLANT. About half the people who experimented with it made no appreciable improvement in yield, although a 16.5% improvement was possible. It is conceivable that we often miss making improvements in real problems by using a weak research strategy.

Researchers using the designed experiment approach on the SNAFU Unit did better than those using the one-test-at-a-time approach. However, insufficient data were collected on this unit to be certain that this would always hold.

The Box-Wilson technique didn't show up to advantage over the designed experiment strategy in either ALCOHOL PLANT model. This may be due to a lack of experience on the part of the researchers, or to a weakness in the technique in problems with large experimental errors.

Not enough people tried other techniques such as the random elimination of operating space or the concept of randomizing interaction vectors to give these strategies a fair try-out.

Further work is needed using well-defined strategies to obtain a good comparison between different strategies. Where the strategy can be delineated in detail, it might be possible to set the problem up on a computer. Thousands of trials with a certain strategy could then be made until a firm value is obtained for the "expected return" for that particular strategy. The difficulty at present is that we do not know how to delineate the strategy in sufficient detail to do this. For example, no firm criterion has been set up for when to quit doing research work. In these games this was left to the judgment of the research workers. The work done so far suggests that the use of designed experiments by experienced people can beat the orthodox one-variable-at-a-time strategy used by experienced people.

V. GENERAL OBSERVATIONS ON HOW PEOPLE DO RESEARCH

A. Luck, Skill, and Judgment

Three elements seem to be important in determining the degree of success. They are luck, skill, and judgment. Those doing well had a more-than-average share of at least one of these elements. Some people were lucky during their first runs in the program, that is, they struck high yield values almost immediately. Where they had the good judgment to quit research at this point, they made a sizable profit. The statisticians and mathematicians tended to depend less on luck and more on skill. The element of judgment, however, was still important. For example, some statisticians made the mistake of covering too small a part of operating space. Some of them did not know when to quit research work.

B. Emotions Aroused

The same emotions seemed to be aroused in the different research teams. Usually, the team started the game with great interest. This was followed frequently by a feeling of jubilation when they obtained a relatively good yield response. However, in many cases this was again followed by frustration when they discovered that they were

unable to check this yield value or to improve on it.

C. The Mistakes We Make

Work with these models has pointed up some of the errors we are apt to make in doing our research work.

1. We aren't bold enough. For example, many research teams never looked at half of the range on one or more of the variables. The best value might lie in the range not studied. Other teams started off with steps which were too small. They investigated a small range around the starting point rather than first making a sketchy study of the entire operable space. An element of boldness seems to be important in doing research. Most of the management people had this.

2. We are apt to get in a rut. Two or three good results in a row will convince most of us that we have a good point. Sometimes we are so convinced that we refuse to look at other points. Unfortunately, a large positive error can occur two or three times in a row. This leads to difficulty. Once the researcher becomes convinced that the best value for a certain variable is in a certain range, he may do a tremendous amount of research before changing his mind. One team ran three tests and concluded that the best level on a variable was 65. They then ran 24 tests without looking at any other level on that variable. The best level was actually at 18! We are apt to get a notion and stick to it through thick or thin.

3. We don't pay enough attention to the error. This leads to frustration. We are trying to attribute every change in yield response to a change we made in one or more of the variables. We don't have a good "feel" for what error can do to our results.

4. Some of us have a linear complex. Intuitively we think of all effects as straight line relationships. This can lead to mistakes.

5. We don't always eliminate the unimportant variables. In the first model of the ALCOHOL PLANT, variable D had no effect on the system. Many of the researchers never discovered this. Elimination of the unimportant variables simplifies the problem.

6. We don't know when to quit. Some research teams failed to stop and consider periodically whether they were really getting anywhere. Some sort of running estimate of how we are doing seems to be essential to a good strategy. The administrators in general had this ability. They quit research early if they discovered that they were not getting anywhere.

VI. ELEMENTS NEEDED IN A "BEST" STRATEGY

The work reported in this paper shows how we do our research work. It points up some of the mistakes we are apt to make but does not reach a final conclusion as to the best research strategy. The designed experiment strategy is better than the one-at-a-time strategy. However, experience and judgment are important. The use of a designed experiment alone does not produce good results. Judgment is needed in making decisions. The designed experiment is probably not the best strategy.

It is simply a better strategy than the orthodox one-variable-at-a-time way of doing empirical research.

Observation of the various strategies used has pointed up some elements which may be part of the "best" strategy.

1. A best strategy will have some way of making use of information as it is collected. It does not seem sensible to wait until a large block of information is collected before changing our strategy or before the information can be used. In the best strategy, decisions will be made more frequently than they normally are when using a standard statistical design. The one-test-at-a-time approach is an unsuccessful attempt to do this.

2. The best strategy will involve a scheme, whereby, the strategy is developed as information is obtained, in short, the master strategy will involve changing the immediate strategy to fit the problem. Changes in the immediate strategy will be made as the nature of the problem becomes more evident. The immediate strategy will then always be the best one for the problem as it unfolds.

3. The best strategy will involve different levels of screening. There will be some gross type of examination first which is made quickly and cheaply followed by a more refined study on a smaller part of operating space and this in turn will be followed by a still more intensive study on a smaller part of operating space. Unimportant variables will be dropped early.

4. The best research strategy must include some way of defining operating space. In most real problems this is difficult. In the model problems used here, of course, operating space was defined automatically by the limits on the dials, there were no islands or inlets in operating space. However, in real problems the operating range on each variable may depend on where the other variables are set.

5. The best strategy will incorporate a boldness which will make it willing to look into the unknown, so that all parts of operating space are explored.

6. Decisions will be made with a calculated risk. The risk taken will probably change as the work progresses. Possibly more risk will be taken in the early stages of the problem than in the final stages.

7. The best strategy will include a rule for deciding when to quit research. It will show us how to know when further research is unlikely to pay off.

VII. SUMMARY

Some research on empirical research strategies has been carried out using electrical analogs of research problems. This work was limited largely to learning how people do empirical research work and to comparing the success of those people who use the one-test-at-a-time approach with those using some type of statistically designed experiment. The statistically designed experiment is better than the one-variable-at-a-time strategy. The work has shown some of the mistakes

we are apt to make in doing empirical research and has suggested some of the elements needed in a best strategy. The help of mathematicians in developing a best strategy is sought. A best strategy for doing empirical research work would be of value in making industrial research more useful to society.

VIII. ACKNOWLEDGMENTS

The work discussed in this paper represents the efforts of many others in the Esso Research and Engineering Company. The authors thank them for their help.

TABLE I

FIRST ALCOHOL PLANT

(Simple Relationships - Unknown Large Error.)

Game Ranking	Strategy Used	Number of Tests	Gain in Yield, %	Profit over Competitive Research
-	Maximum Possible Gain	-	18.2	-
1	One variable at a time, duplicated first three tests. Luck? Know when to quit.	10	16.6	$7,100,000
2	One at a time, large steps (10-90). Didn't cover all variables. Luck? Know when to quit.	9	15.6	6,720,000
3	1/4 replicate of 2^5 factorial (35-65) then small adjustments.	19	17.2	6,420,000
4	Random numbers, then small adjustments.	18	14.2	4,940,000
5	Randomize interaction vectors.	17	13.9	4,910,000
6	One at a time, small steps at first. Regression analysis.	18	13.0	4,340,000
7	Box-Wilson starting with 1/4 replicate of 2^5 factorial (35-65).	33	16.3	4,190,000
8	Luck. Know when to quit.	3	8.0	3,640,000
9	One variable at a time, replicate to get estimate of error.	18	10.6	3,140,000
10	1/4 replicate of 2^5 factorial (40-60).	10	7.0	2,300,000
11	One test at a time. Know when to quit.	6	4.8	1,680,000
12	1/4 replicate of 2^5 factorial (40-60).	11	5.8	1,580,000
13	One at a time (25-75). Know when to quit.	11	5.7	1,530,000
14	1/4 replicate of 2^5 factorial (35-65).	13	5.7	1,290,000
15	One at a time. Became frustrated.	18	6.8	1,240,000

TABLE I (Cont'd.)

FIRST ALCOHOL PLANT

Game Ranking	Strategy Used	Number of Tests	Gain in Yield, %	Profit over Competitive Research
16	One at a time. Became frustrated.	39	11.5	$1,070,000
17	One at a time. Made many small changes.	18	6.0	840,000
18	One at a time. Became frustrated.	19	6.2	820,000
19	One at a time. Frustrated at end.	23	5.7	90,000
20	1/4 replicate of 2^5 factorial. Small steps. (50-65)	17	4.1	10,000
21	One at a time. Frustrated at end.	25	5.0	-500,000
22	Full 2^5 factorial design (20-80). Data not analyzed.	43	8.5	-910,000
23	1/4 replicate of 2^5 factorial. Small steps (50-60). Then one at a time. Frustrated.	23	3.1	-1,210,000

TABLE II-1

SECOND ALCOHOL PLANT

(More Complex Relationships - Error Smaller and Known.)

Game Ranking	Strategy Used	Number of Tests	Gain in Yield, %	Profit over Competitive Research
-	Maximum Possible Gain	-	16.5	-
1	Fractional replicates of 2^5 factorial.	20	15.9	$5,550,000
2	1/9 replicate of 3^5 factorial.	31	16.3	4,430,000
3	Eliminate operating space using random numbers.	9	8.5	3,170,000
4	One variable at a time, big steps (10-80) Luck?	9	8.5	3,170,000
5	1/4 replicate of 2^5 factorial, big steps (20-75).	15	9.1	2,750,000
6	Randomize interaction vectors.	18	9.1	2,390,000
7	1/4 replicate of 2^5 factorial repeated at different levels (20-60 and 40-80).	20	9.5	2,350,000
8	Series of fractional 2^n factorials.	33	12.0	2,040,000
9	Small designed experiments.	27	7.8	660,000
10	Box-Wilson Technique.	34	9.0	420,000
11	Common sense, intuition, one variable at a time.	33	8.2	140,000
12	1/4 replicate of 2^5 factorial repeated.	15	3.5	-50,000
13	1/4 replicate of 2^5 factorial.	9	2.0	-80,000
14	One at a time, big steps (15-75).	21	4.5	-270,000
15	One at a time, big steps (20-80).	13	5.0	-500,000
16	Factorial designed experiments (20-70).	46	9.8	-620,000
17	Small steps, one at a time (45-60).	8	0	-960,000
18	One at a time.	8	0	-960,000
19	One at a time.	8	0	-960,000
20	One at a time, big steps (10-90).	21	3.0	-1,020,000
21	One at a time. (40-70)	9	0	-1,080,000

TABLE II-2

Game Ranking	Strategy Used	Number of Tests	Gain in Yield, %	Profit over Competitive Research
22	One at a time. (25-70)	9	0	-1,080,000
23	One at a time. (30-70)	19	1.6	-1,480,000
24	One at a time. (50-70)	14	0	-1,680,000
25	Two at a time. (25-70)	14	0	-1,680,000
26	1/4 replicate 2^5 factorial (35-65). Then Box-Wilson.	25	2.4	-1,800,000
27	One at a time.	15	0	-1,800,000
28	One test at a time, big steps (5-85).	23	1.0	-2,260,000
29	One test at a time. (25-75)	22	0	-2,640,000
30	1/4 replicate of 2^5 factorial (35-65), plus other tests.	27	1.1	-2,690,000
31	Latin Square type of design (5 x 5). Small steps (40-60).	28	1.0	-2,800,000
32	One at a time (25-70).	44	1.1	-4,730,000

TABLE III

SNAFU UNIT

(With Two Discrete Variables - Small Error Known.)
(Each Team Was Allowed 40 Tests.)

Game Ranking	Strategy Used	Yield Improvement	% of Possible Gain
-	Maximum Possible Gain	14.0%	100
1	27 tests in form of 1/27 fraction of 3^6 factorial, then 1/2 replicate of 2^4 factorial on 4 continuous variables.	12.7%	91
2	Fractional replicates of 2^n factorial experiments.	11.5%	82
3	One test at a time.	10.3%	74
4	One test at a time	8.6%	62
5	One at a time or several at a time.	6.3%	45
6	One at a time.	3.7%	26

STATISTICAL METHODS IN SORTING OPERATIONS

William C. Vissing
Eli Lilly and Company

Wherever it is possible, we attempt to prevent the occurrence of defective pieces. In some processes it is only economically feasible to keep defectives to a minimum. We resort to sorting in these operations to further reduce the defectives within the limits of our outgoing quality levels.

The purpose of this paper is to explain how sampling and charts can be applied effectively to maintain quality and reduce labor and material costs in sorting operations. There are additional benefits. Three different sorting operations will be explained to show the application of techniques and specific benefits derived in each program.

I.
Soft Gelatin Capsules

In the gelseal manufacturing department the product was given a 100 percent sort after the drying operation, for the purpose of removing the defective gelseals (soft gelatin capsules). The most prominent defective was the leaker. This was a capsule that lost cottonseed oil during the drying operation. Very often oil spots were left on the absorbent paper used on the drying trays. This made it relatively easy to see the defective and remove it. Others have seeped oil only slightly and could not be seen. These could become a serious problem after extended periods of storage.

Our pilot plant developed a method to detect leakers more readily. Since some of the vitamins are fluorescent, they can be seen easily under a harmless ultra-violet light. It became relatively easy to spot the defective piece and remove it.

Other defectives of less intensity were malformed capsules, specks, bubbles, and partial fills. The trays were examined very carefully for all these defectives and then emptied into a cardboard box. These boxes (containing approximately 15,000) were examined by an inspector for the "pharmaceutical elegance" of the product.

These questions arose: What is our present quality level? Is it necessary to do so much sorting? Would it be possible to remove only the leakers or obvious defectives by a quick sort and then, by the use of a sample, determine whether further sorting is necessary?

Working very closely with the development and production divisions, we established a set of standards that were acceptable to the control division. Some of the production was sampled for approximately a month, after the

regular sorting operation, to determine the existing quality level. Using this as a guide, an acceptable quality level was set for the process.

We agreed to compare the quality of lots of material after conventional sorting (a detailed examination of each tray) with a new method (a quick inspection of each tray to remove obvious defectives). It was determined that if we found and removed seven (7) or less obvious or critical defectives on a tray after a quick inspection this tray would be emptied into a box with no further sorting. If eight (8) or more were found and removed, that tray was given a detailed inspection for all defectives. At the end of this process a sample of three hundred (300) was removed from each box and examined carefully for all defectives. The plan was to accept zero (0) critical defectives and seven (7) or less obvious (minor) defectives. Resorting was performed on those that did not pass. The results of this experiment were as follows:

	Total Defect. Before Sorting	Defect. Removed In Sorting	Defect. Remain. After Sorting	Total Boxes Exam.	Boxes 100% Sorted
8 Lots Old Method	1.09%	.58%	.51%	221	221
11 Lots New Method	1.08%	.42%	.66%	304	12

The AQL of the lots sorted by the new method was approximately the same as the AQL of the conventionally sorted lots. Intensive sorting was devoted to those sections that needed it, while a quick sort was all that was necessary for the major portion of the eleven lots.

This program was adopted essentially as described above. The methods department determined that there was a monetary savings of approximately $44,000 annually. The program was established on a trial basis and eventually became a part of the regular procedure for this department.

Where the sample indicated it was necessary to sort, very often one to two types of defectives were predominant. The sorter could concentrate her efforts on these and thus do a more effective job. Manufacturing and development personnel were very quick to note high amounts of certain types of scag (defective) and make process improvements. This was an additional benefit in the program.

After approximately a year of operation, the number of boxes being 100 percent sorted dropped from 7 percent to less than 1 percent. This was attributed to further improvements in the quality by the pilot plant and production departments. We eventually progressed to reduced sampling where one out of every four boxes was sampled.

The overall benefits of this program have been:

1. Uniform Quality Level
2. Increased Yield
3. Reduced Labor Costs
4. Improved Sorting Techniques

II.
Timed Disintegrating Tablets

There is a definite need to establish sorting standards on a product. First of all, it defines a defective and also an acceptable piece. It assists in training new sorters, serves as a review for experienced sorters and helps to keep the number of good pieces removed as defective to a minimum. This last point is one of the human errors in sorting; that is, removing good pieces as defective. Sorting can be made more effective when there is concentration on a particular defective or group of defectives with less emphasis on the minor categories.

This was important in the sorting operation in the process of manufacturing timed disintegrating tablets. Tablets pass before two sorters. The first sorter looks at one side of the tablets and removes defective pieces. The tablets are automatically inverted and another sorter observes the other side of the tablets. One of the first analyses of this operation showed there were 35 to 50 percent acceptable pieces in the scag. Obviously, standards were a must.

The categories on standards were as follows:

Critical - Any type of defective that impairs the value of the tablets (broken, double tablets).

Major - An obvious defective that renders the tablet unsightly, but does not affect the therapeutic properties of the tablet (large bubbles, large scratches).

Minor - A defective not defined as major or critical (small scratches).

The standards were explained to the sorters. They were asked to concentrate their efforts on the major and critical categories with less attention to the minor or border line tablets. The results were that acceptable tablets removed as defective dropped from 35 to 50 percent to approximately 6 percent. The resultant monetary savings were substantial. By concentrating on the critical and majors, there was also an improvement in the outgoing quality level.

Two charts were adapted to the process. One was to control the percent of good removed as defective. We established an arbitrary upper limit of 10 percent good

tablets in the product removed as defective scag. Samples of three hundred (300) were taken from approximately every fifth lot by the supervisor and analyzed for the percent good in the scag. These values were plotted on a chart. This became a very useful tool in training new sorters.

The other chart was a "c" chart on scag removed in sorting. The average scag of five (5) consecutive lots was plotted on the chart. C varied from .10 to .30 of one percent. The upper limit was used to warn the process operators that they were producing 'enseals' at a high defective rate.

Effective control on the quality of the sorted product was maintained by an inspection department.

'enseals' (Timed Disintegrating Tablets, Lilly)

III.

Sorting of Filled Capsules

The sorting is combined with dusting and polishing pulvules (filled capsules). As the operators dust and polish the product on large tables, they stop intermittently to examine the pulvules for defectives. The acceptable product is put into drums, checked by inspection for "pharmaceutical elegance" and then sent to the finishing department.

Good quality is important in subsequent operations for two major reasons.

1. To keep down-time on automatic filling equipment to a minimum.

2. To reduce rework of the product after packaging.

Critical defectives, such as separated capsules, made it necessary to stop and clean equipment frequently. Operators took off bottles that contained visible defectives as they passed before them on the finishing belt. Both factors increased the cost.

What should be the outgoing quality level? What is it now? There were no numerical answers to these questions. Standards of "pass this or better" nature were established. Then sampling was applied to determine the AQL.

The pulvules fell into two groups; old process and new process, and therefore two quality levels. (The company was in the process of shifting all capsule manufacturing to a new plant.) Quality levels were 3 percent and 1.5 percent respective with normal sorting operations. These values were incorporated into acceptance sampling plans.

The procedure, very briefly, is as follows:

Two production inspectors sample sorted drums of pulvules. Those found to be unacceptable by the sample are returned for resorting on the sorter's own time. The sorter is informed of the type of defectives most prominent in the sample. This provides for more selective resorting and improved performance.

The statistical inspection department maintained charts for a time in the department to show the AQL of each lot in the two groups.

Sampling is applied also to the scag. Results on the amount of good in the scag are reported to production supervision. The statistical inspection department samples two or three sections of scag per week as an audit of the process.

The results of the program have helped reduce reworks and down time, and also improved the outgoing quality.

Supervision became aware of certain types of defectives made at the filling machine. They organized for corrective action. A program was initiated to periodically inspect filling rings for excessive surface wear. Worn rings were replaced. This further helped to reduce scag and improve the outgoing quality.

Summary

We strive for perfection in detailing but have come to recognize there are definite limits in these operations. You can do much to assist the sorter to improve the operation; such as, provide adequate equipment and light, define defectives through standards, determine process capabilities, provide for indoctrination of your program and follow up with examinations of the product removed as defectives.

These steps can lead to labor and material savings, but the methods can often illustrate where further improvements can be made in the process. Therein lies one of the basic concepts of quality control; building quality into the product at the machine. As the types of defectives are reduced to a minimum, sorting can be minimized as well as made to be more effective. These will help attain the goal of a higher quality product at minimum cost.

STATISTICAL STUDIES OF ROSIN SIZING EVALUATION VARIABLES:

IV. THE DETERMINATION OF MINIMUM-COST TESTING PROCEDURES

W.M. Wooding
formerly American Cyanamid Co.;
now Revlon, Inc.

Introduction

In a laboratory conducting sizing evaluation work, it quite often becomes desirable to alter the sheetmaking and related testing procedures in the interests of economy, efficiency or reliability. So many factors go to make up the sum total of a given evaluation, however, that it is often difficult to develop a system by which comparisons between procedures may be made without resorting to guesswork as at least a partial basis for the decisions that must be reached.

The methods described in this presentation will allow comparison of two sets of evaluation conditions, such as the use of different sheet machines, different refining methods, etc., as well as the determination of the proper number of tests and sheets to use in a given procedure. Since results may be expressed as costs, either in dollars or in time consumed, they are readily presentable to persons who may not be familiar with the statistical techniques involved. The limitations on the use of such a system as this, of course, lie in the fact that results apply only to the predicated conditions; therefore, if alternative procedures are foreseeable, they must be included in the experimental design. Provided that this can be done, the decision as to the best and most economical procedure becomes simply a matter of comparing costs or time consumed.

The procedures to be described have been adapted to other systems in addition to rosin size evaluation work, and one of the purposes of this paper is to suggest such generalizations. It should therefore be borne in mind that although the examples selected below were based on rosin size work, they are very definitely not limited to this field.

The remainder of this paper will be presented as a report on a project carried out at Cyanamid, which will illustrate these computational procedures.

Experimental Design and Procedure

The questions asked at the beginning of this experiment were:

1) In evaluating rosin sizes, how many handsheets should be made, and how many sizing tests should be made per sheet, for lowest cost operation (cost per sample of rosin size tested)?*
2) Does the use of additional refining following beating, using the Morden Laboratory Refiner, reduce or increase the cost of sizing evaluation?

The information presented below represents a portion of a broader experiment, in which a number of additional variables including

*It was believed, and subsequently shown, that the figures obtained would differ with different types of sizing tests.

degree of beating, type of sheet machine used and other factors were included in order to measure their effect upon the cost of evaluation of a size.

To answer the questions proposed, the following experiments were planned and carried out. All laboratory procedures referred to below have been described in a previous publication (2):

Two rosin sizes were used, adding 1% size solids and 1½% alum (dry basis) to the pulp. The experimental design was in the form of a small factorial:

$M_1Z_0A_0$ $M_1Z_0A_1$
$M_1Z_1A_0$ $M_1Z_1A_1$
$M_0Z_0A_0$ $M_0Z_0A_1$
$M_0Z_1A_0$ $M_0Z_1A_1$

where,
M_0 = Pulp not "Mordened"
M_1 = Pulp "Mordened" for 2 minutes
Z_0 = Regular rosin size
Z_1 = Fortified rosin size
A_0 = Sheets tested on felt side*
A_1 = Sheets tested on wire side*

The sheetmaking work of this experiment was done in a single day, so that no question existed regarding the effect of day-to-day variations in sheetmaking conditions upon the data or their precision.

The handsheets made were given the following sizing tests:

Ink Penetration Test: These were so-called "optical" or modified BKY tests, using an instrument designed at Stamford similar to the regular BKY Ink test, but providing for elimination of the effects of heat from the light source upon the measurements.

Lactic Acid: Standard Penescope test at 100 °F.

Total Water Immersion: "Water Absorption" test, carried out by immersing weighed portions of sheets in water at 70 °F. for 15 minutes, blotting, reweighing and expressing results as per cent gain in weight.

Basis Weight: These data were obtained as a check upon the uniformity of the sheet batches made, and were done in the conventional manner.

The testing schedules used were as follows:

Ink Tests:

Wire Side: Two tests on each of five light sheets**, using sheet Nos. 3, 5, 7, 9, 11 out of the serially-numbered total of 15 sheets made.

Felt Side: Same, on sheet Nos. 2, 4, 6, 8, and 10.

Total, 20 tests per set.

* At the time this work was done, it was known that differences in precision between wire and felt test data existed in some cases (cf. Ref. 1 and 2).

**The sets of handsheets consisted of 15 light sheets (50 lb.) and eight heavy sheets (200 lb.), 25 x 40/500 basis.

Lactic Acid Tests:

Wire Side: Two tests on each of two heavy sheets, Nos. 1 and 5, out of the total of eight sheets made.

Felt Side: Same, using sheets 3 and 7.

Total, Eight tests per set.

Water Absorption Tests:

Two tests on each of four heavy sheets, Nos. 2, 4, 6 and 8.

Total, Eight tests per set.

Basis Weight:

Average figure obtained by weighing groups of sheets from each set.

Summary of Data Obtained

The basis weight data obtained for each set are shown in Table I. The individual test data are given in Tables II, III and IV for Ink, Lactic Acid and Water Absorption tests respectively.

Table I
Summary of Basis Weight Data

Experimental Code	Explanation	Basis Wt., Lb., 25 x 40/500 Light	Heavy
M_1Z_0	Mordened; regular size	49.3	209.0
M_1Z_1	Mordened; fortified size	49.7	208.5
M_0Z_0	Not Mordened; regular size	48.8	205.5
M_0Z_1	Not Mordened; fortified size	48.2	201.0

Computation of Number of Tests Per Sheet and Total Number of Tests

The purpose of the following calculations was to determine the net standard deviation of a single item when samples are examined in various combinations of tests between and within sheets. This is done by analyzing, then recombining the sheet-to-sheet and within-sheet estimates of error. The resulting values of standard deviation are then used in conjunction with Sillitto's tables (3) as described previously (1), but this time as separate determinations of the number of tests required, using various numbers of tests per sheet.

It will be seen that if the within-sheets error is low and the between-sheets error is high, it will be advantageous, in terms of the total number of tests needed, to make few tests per sheet and to use a relatively large number of sheets. If, on the other hand, the within-sheets error is high and the between-sheets error is low, fewer sheets and more tests per sheet will be necessary in order to obtain the smallest number of tests. To illustrate the calculations involved, the ink test data of Table II will be used as an example:

The data were classified into small groups, each representing a single level of the variables concerned. Assuming that the effect of Mordening upon precision is to be determined, but that of wire versus felt and that of different sizes is not*, the classifications would be

*These had been measured in previous experiments.

TABLE II

Ink Penetration Test Data
(light sheets)

Explanation	Exp. Code	Wire Side Tests, Seconds									
		Sheet 3		Sheet 5		Sheet 7		Sheet 9		Sheet 11	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Mordened,* Regular Size	M_1Z_0	215	215	210	215	215	220	205	240	265	265
Mordened, Fortified Size	M_1Z_1	230	225	230	250	225	260	235	250	235	230
Not Mordened, Regular Size	M_0Z_0	255	285	270	280	265	285	250	275	260	270
Not Mordened, Fortified Size	M_0Z_1	260	260	285	280	310	300	295	265	295	305

		Felt Side Tests, Seconds									
		Sheet 2		Sheet 4		Sheet 6		Sheet 8		Sheet 10	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Mordened, Regular Size	M_1Z_0	205	190	210	215	220	225	235	200	200	190
Mordened, Fortified Size	M_1Z_1	190	215	200	220	200	200	210	210	200	230
Not Mordened, Regular Size	M_0Z_0	280	295	270	255	290	285	260	310	270	245
Not Mordened, Fortified Size	M_0Z_1	295	255	260	245	280	290	260	300	285	285

* When Morden was used, stock had already been sized. This may account for the reduction is sizing with Mordening. However, it will be noted that Mordening apparently improved Lactic acid tests.

TABLE III

Lactic Acid* Test Data
(heavy sheets)

Explanation	Exp. Code	Wire Side, Seconds				Felt Side, Seconds			
		Sheet 1		Sheet 5		Sheet 3		Sheet 7	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Mordened, Regular Size	M_1Z_o	120	130	100	115	120	125	65	110
Mordened, Fortified Size	M_1Z_1	195	185	230	175	155	170	165	215
Not Mordened, Regular Size	M_oZ_o	85	70	55	60	50	55	50	55
Not Mordened, Fortified Size	M_oZ_1	90	115	75	85	80	80	70	75

* 15 minutes' immersion at 70° F.

TABLE IV

Water Absorption* Test Data
(heavy sheets)

Explanation	Exp. Code	% Absorption							
		Sheet 2		Sheet 4		Sheet 6		Sheet 8	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Mordened, Regular Size	M_1Z_o	44.0	39.3	41.0	43.6	40.7	41.7	40.8	41.9
Mordened, Fortified Size	M_1Z_1	35.3	36.8	36.4	35.0	36.8	35.0	35.3	35.3
Not Mordened, Regular Size	M_oZ_o	36.6	36.5	39.6	39.6	38.3	38.3	38.1	36.3
Not Mordened, Fortified Size	M_oZ_1	32.7	32.1	35.1	32.4	35.1	34.5	33.6	34.3

* 100° F.

the following:

Mordened	Not Mordened
M_1Z_0 Wire	M_0Z_0 Wire
M_1Z_1 Wire	M_0Z_1 Wire
M_1Z_0 Felt	M_0Z_0 Felt
M_1Z_1 Felt	M_0Z_1 Felt

An analysis of variance was carried out upon each small group, bearing in mind that ink tests were done at the rate of two per sheet, using five sheets for the felt and five for the wire side tests. The degrees of freedom (DF) were summarized as follows: (Table V)

Table V
Analysis of Variance
(Example--M_1Z_0 Wire)

Source of variation	DF
Between sheets (5 sheets)	4
Within sheets (1 DF/sheet, 5 sheets)	5
	9

Sums of squares were calculated conventionally to outline an analysis of variance as shown for each of the classifications used. The following formulas were used in this particular work, and represent simplified computations, since only two tests per sheet were used: *

1) $SS_b = \frac{\Sigma(X_1 + X_2)^2}{2} - \frac{(\Sigma X)^2}{10}$

2) $SS_w = \frac{\Sigma(X_1 - X_2)^2}{2}$

3) $SS_t = \Sigma X^2 - \frac{(\Sigma X)^2}{10}$

where X = An individual measurement
X_1 and X_2 = A pair of measurements on the same sheet
SS_b = Sum of squares for "between sheets"
SS_w = Sum of squares for "within sheets"
SS_t = Total sum of squares

The use of the SS_t computations is a check upon the calculations of the other two.

The analyses of variance were tabulated and the sums of squares added (pooled) with the results shown in Table VI (the "unmordened" data were handled in the same way).

* Reference (4) gives generalized formulas for these operations.

Table VI
Pooled Sums of Squares
(Stock Mordened)

Set	SS_b	SS_w	SS_t	DF_b	DF_w
M_1Z_0 Wire	3815	638	4453	4	5
M_1Z_0 Felt	1190	800	1990	4	5
M_1Z_1 Wire	360	950	1310	4	5
M_1Z_1 Felt	300	963	1263	4	5
	5665	3351	9016	16	20

The pooled values of SS_b and SS_w were divided by their respective degrees of freedom, to obtain mean square values for between and within sheets (MS_b and MS_w):

1) $MS_b = \frac{SS_b}{DF_b} = \frac{5665}{16} = 354$

2) $MS_w = \frac{SS_w}{DF_w} = \frac{3351}{20} = 168$

now, by the law of additivity of variances,

$$\sigma_b^2 + k\sigma_b^2 = (EMS)_b$$
$$\sigma_w^2 = (EMS)_w$$

where EMS = "expected mean square"
k = actual number of tests per sheet (2)
therefore, (substituting):

$$2\sigma_b^2 = (EMS)_b - (EMS)_w$$
$$\sigma_b^2 = \frac{(EMS)_b - (EMS)_w}{2}$$

and, in terms of the estimated values of σ^2 (which are S^2), and EMS (which are MS):

1) $S_b^2 = \frac{(MS)_b - (MS)_w}{2}$

2) $S_w^2 = (MS)_w$

Substituting the values found for $(MS)_b$ and $(MS)_w$,

$$S_b^2 = \frac{(MS)_b - (MS)_w}{2} = \frac{354 - 168}{2} = 93$$

$$S_w^2 = (MS)_w = 168$$

NOTE: Since the within sheets MS, $(MS)_w$, must obviously always be equal to or smaller than that for between sheets, $(MS)_b$, which includes it, we must conclude that $(MS)_b$ and S_b^2 are both equal to zero whenever it is found experimentally that $(MS)_w$ is greater than $(MS)_b$. This situation, of course, did not exist in this particular case, and it is obvious that there is considerable sheet-to-sheet error in addition to the within sheet errors; i.e., $(MS)_b$ is greater than $(MS)_w$.

Variances and standard deviations to be used in computations of the total number of tests by entering Sillitto's tables were now calculated by means of a relationship derived as follows:

Let A = number of sheets tested

B = number of tests per sheet

$S_{\bar{x}}^2$ = variance of means

Now, by additivity of variances,

$$ABS_{\bar{x}}^2 = S_w^2 + BS_b^2 \qquad (AB = \text{total number of tests})$$

By definition,

$$S_{\bar{x}}^2 = \frac{S^2}{AB}$$

Therefore, substituting,

$$S^2 = S_w^2 + BS_b^2$$

Using this relationship, several values of B were selected arbitrarily and S^2 calculated for each:

Recall that $S_b^2 = 93$ and $S_w^2 = 168$.
B = 1, 2, 3 and 4 * tests per sheet

The calculated results are shown on Table VII:

Table VII
Calculated Overall Standard Deviations
For Various Numbers of Tests/Sheet
($S^2 = 168 + 93B$)

B =	1	2	3	4
S^2 =	261	354	447	540
$S = \sqrt{S^2}$ =	16.16	18.81	21.14	23.24

* 4 was the maximum usable value because of the size of the British handsheets.

The calculated values of S were used to enter Sillitto's tables (3) as described in a previous publication (1), using selected values of CD *. For this work, the α and β errors chosen were 5 and 10% respectively. To facilitate this portion of the work, plots of Sillitto's tables were used. Reproductions of these have been published previously (1). It should be noted that Sillitto's tables are based upon the t-test, although they have the additional advantage of allowing an estimate of error. For this reason, values for the number of tests required will be low for a given confidence level, when it is intended to compare pairs of means selected from groups containing more than two means. For such work, a more realistic comparison of means is furnished by the Tukey method. However, for the purpose of this investigation, which was to obtain a relative evaluation of numbers of tests and subsequently of costs, the absolute values were not necessary.

Using the method just described, plots were prepared, showing the number of sheets and tests needed for various number of tests per sheet and desired critical differences. These are shown for the three tests, ink, lactic, and water, for each of the two conditions, Mordened and not Mordened, in Figures 1, 2 and 3. Intermediate curves for 2 and 3 tests per sheet are omitted for clarity.

It will be noted that the total number of sheets and tests, and therefore the total cost of an evaluation, goes up very sharply as the desired CD is reduced in arithmetic progression. This points up the necessity of carrying out evaluations at the highest CD value which can be used, consistent with the type of information desired.

The preceding data were next applied to calculations of the relative costs per size sample evaluated, in order to make possible a decision as to the optimum number of tests and sheets for standard laboratory work when these three types of tests were used to evaluate a single size.

CALCULATIONS OF EVALUATION COSTS

Computations of the number of sheets and tests which could be used to evaluate a given size, as has been illustrated above, does not complete the picture. For most economical operation, it remains to choose the combination of sheets and tests representing the lowest cost, and this, of course, is dependent upon the relative expense of sheetmaking and testing. The necessary calculations to allow such decisions are illustrated below.

In performing the following calculations, the assumption was made that, with one exception, the cost of materials used in sizing evaluations was negligible relative to the cost of time consumed. The times used for each operation shown were necessarily estimated average values, but as much care as possible was taken to obtain them. The estimated times were converted to dollar values on the basis of arbitrarily-selected overhead cost figures for sheet testing and sheetmaking operations.

* CD = Critical Difference; i.e., least significant difference between two means.

FIGURE 1

INK TEST

NUMBER OF SHEETS PER SAMPLE

VS.

PRECISION OBTAINABLE

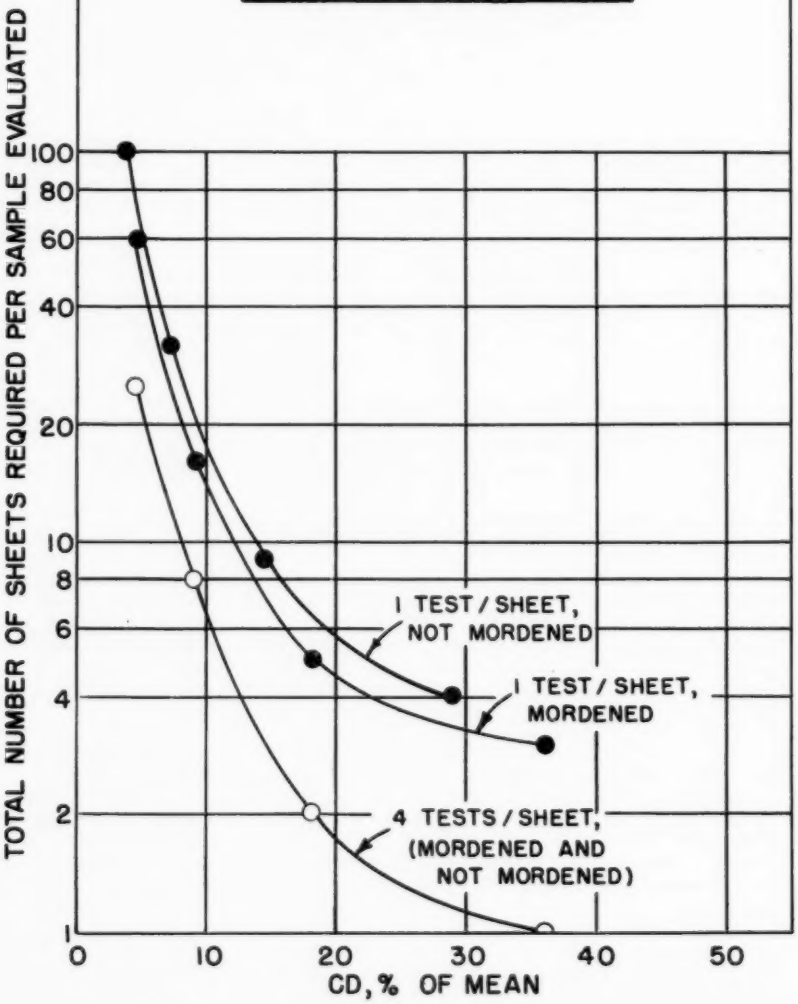


FIGURE 2

LACTIC ACID TEST

NUMBER OF SHEETS vs. PRECISION OBTAINABLE

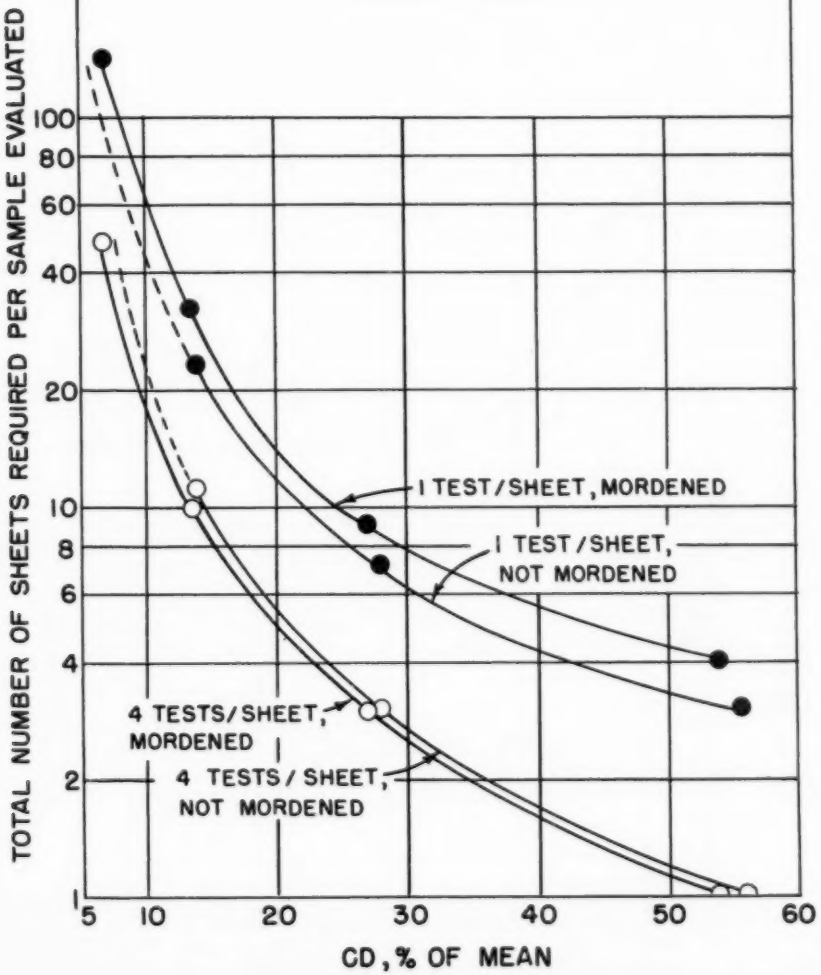
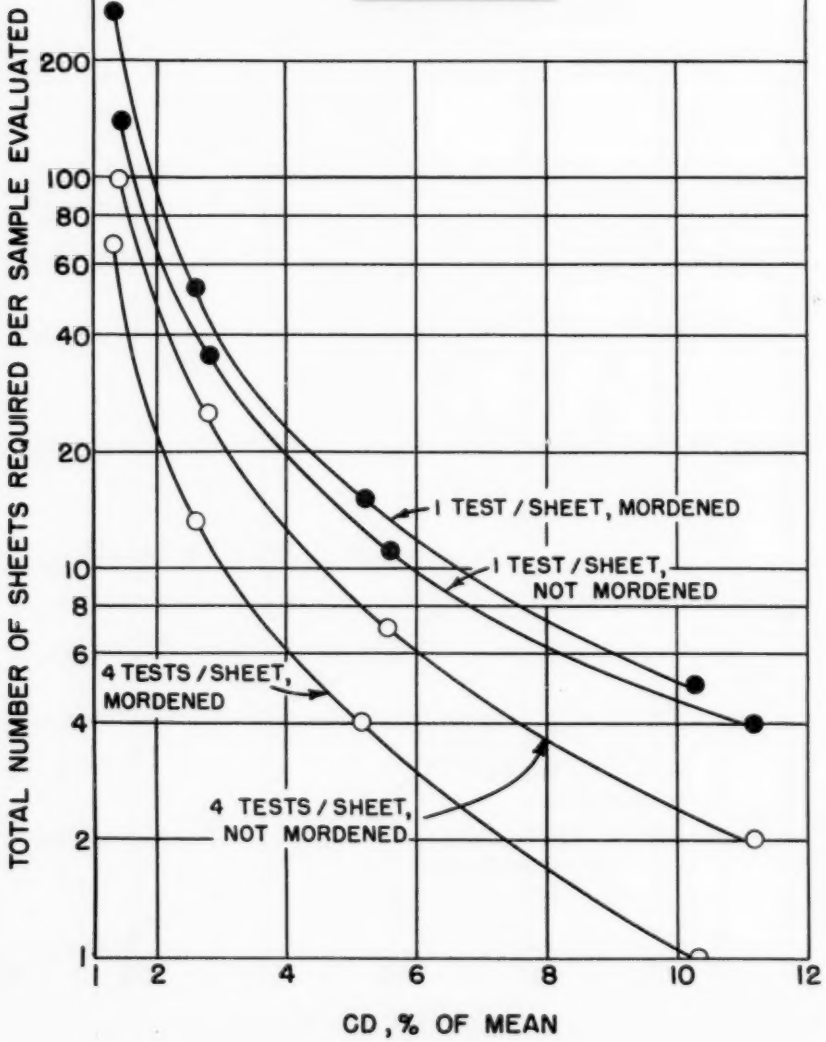


FIGURE 3

WATER ABSORPTION TEST

NUMBER OF SHEETS vs. PRECISION OBTAINABLE



All results were expressed finally in terms of relative dollar costs of a single rosin size, using ink, water and lactic tests. The method illustrated may be extrapolated by analogy to any set of local sheetmaking and testing overhead costs, and any testing schedule.

At Cyanamid, using the present laboratory procedures, the cost of evaluating a rosin size sample consists primarily of the following major components: costs of beating, Morden refining, blotters, handsheet making and testing. These are analyzed below using the assumed overhead costs:

a) Beating: A regular beater run with the "5 pound" Valley beater at Cyanamid contains approximately three thousand grams of pulp (dry basis). Three such runs may be made by one man in a day. Therefore nine thousand grams of pulp are available after a day's work. Assuming an overhead charge of $6.00 per man-hour, this amounts to a cost of $48.00 per nine thousand grams of pulp, exclusive of materials.

"Light" British handsheets weigh about 1.2 grams each. Therefore, 9000/1.2 = 7500 light sheets per day's output. The cost per sheet is therefore $48.00 per 7500 = $0.006 per light sheet. Similarly, "heavy" British sheets weigh about 5.0 grams each. Performing a similar calculation,

$$\frac{\$48.00}{9000/5} - \$0.027 \text{ per heavy sheet.}$$

b) Morden Refining: After sizing, the stock is refined further by treatment in the laboratory Morden refiner (2). About five minutes of total working time are required of a man to "Morden" a batch of stock from which fifteen light (1.2 gram) and eight heavy (5.0 gram) sheets will be made, allowing for discarded test sheets. This requires 95 grams of pulp (dry basis).

Of the 95 grams of pulp Mordened, 58 grams are used for actual sheets to be tested (15 x 1.2 + 8 x 5.0). Mordening time is therefore 5 minutes per 58 grams = 0.0862 minutes per gram of pulp Mordened. At the assumed overhead rate of $6.00 per man-hour for sheetmaking, the cost becomes:

$$0.0862 \times \frac{\text{min.}}{\text{gm.}} \times \frac{\$6.00}{60 \text{ Min.}} = \$0.00862 \text{ per gram.}$$

Therefore, Mordening costs are:

$$\text{Light Sheets: } 1.2 \times \frac{\text{gm.}}{\text{sheet}} \times \frac{\$0.00862}{\text{gm.}} = \$0.010/\text{sheet}$$

$$\text{Heavy Sheets: } \frac{5.0 \text{ gm.}}{\text{sheet}} \times \frac{\$0.00862}{\text{gm.}} = \$0.043/\text{sheet}$$

c) Sheetmaking: The sheetmaking for this experiment required one man-day for four sets, including preparation, calculations and note taking. This is equivalent to two hours per set of sheets. It was observed that about the same amount of time was consumed in making the eight heavy sheets as the fifteen light sheets in each set. Therefore:

Light Sheets: $\frac{60 \text{ min.}}{15 \text{ sheets}} \times \frac{\$6.00}{60 \text{ min.}} = \$0.400/\text{light sheet}$

Heavy Sheets: $\frac{60 \text{ min.}}{8 \text{ sheets}} \times \frac{\$6.00}{60 \text{ min.}} = \$0.750/\text{heavy sheet}$

d) Blotters: The only significant material cost in the evaluation work appeared to be that of blotters. At the time this work was done, 8" x 8" blotters, pre-cut, were available at $6.04/thousand. Every British sheet required three new blotters in its preparation; therefore:

$$\text{Blotter Cost} = \frac{\$0.00604}{\text{blotter}} \times 3 = \$0.018/\text{sheet}$$

e) Testing: A study of the testing work gave the following estimates of testing time:

Ink Tests:	50/man-day
Lactic Tests:	100/man-day
Water Tests:	150/man-day

Assuming an overhead charge of $5.00 per man-hour for testing or $5.00/60 = $0.083/minute:

Ink Tests: $\frac{8 \text{ hr.}}{50 \text{ tests}} \times \frac{60 \text{ min.}}{\text{hour}} \times \frac{\$0.083}{\text{min.}} = \$0.800/\text{test}$

Lactic Tests: $\frac{8 \text{ hr.}}{100 \text{ tests}} \times \frac{60 \text{ min.}}{\text{hour}} \times \frac{\$0.083}{\text{min.}} = \$0.400/\text{test}$

Water Tests: $\frac{8 \text{ hr.}}{150 \text{ tests}} \times \frac{60 \text{ min.}}{\text{hour}} \times \frac{\$0.083}{\text{min.}} = \$0.266/\text{test}$

f) Summary of Costs: The foregoing calculations are summarized, for convenience, in Table VIII.

Table VIII

Costs and Time Consumption for Rosin Size Evaluation Work-Summary
(British Sheetmaking Equipment)

Assumed Pulping, Mordening and Sheetmaking overhead = $6.00/man-hour
Assumed Testing overhead = $5.00/man-hour

Operation	Cost, Dollars per light sheet	Cost, Dollars per heavy sheet	Time, Minutes light sheet	Time, Minutes hvy. sht.
(a) Beating	$0.006	$0.027	0.0640	0.2667
(b) Mordening	0.010	0.043	0.1034	0.4310
(c) Sheetmaking	0.400	0.750	5.0000	7.5000
(d) Blotters (material cost)	0.018	0.018	------	------
(e) Testing				
Ink, per test	$0.800		9.6000	
Lactic, per test	0.400		4.8000	
Water, per test	0.266		3.2000	

Actual Cost Calculations

The calculations to be described below were done on the assumption that a single sample of rosin size was to be evaluated, using ink, lactic acid and water absorption tests only. Costs

rather than time-consumption were used (although either basis would have been valid), because the costs, although relative, seemed easier to visualize and use.

It was assumed that fresh sheets were used for each type of test; i.e., even when only one test per sheet was used, the remainder of the set was not used for other types of tests. This procedure is the only one which is statistically valid when correlation may exist among the various tests used.

Evaluation costs were obtained by simply multiplying the cost per sheet or per test for each operation by the number of sheets or tests required, and tabulating and adding the results. A set of tables corresponding to Figures 1, 2 and 3 were prepared, in which the total cost of evaluation under each condition was indicated as a sum of costs of the individual operations required. In preparing these tables, it was assumed that only the sheets required for actual testing would be made. This, of course, was not necessarily the exact truth, and there would be additional charges if standard sets containing a fixed number of sheets are always made. This was not expected to have any major effect upon the relative costs of evaluation, however.

The total cost was obtained by the use of the following simple formula:

$$C = aP + aM + aS + aB + bT$$

where
- a = number of sheets
- b = number of tests
- P = pulping (beating) cost, dollars
- M = Mordening cost, dollars
- S = Sheetmaking cost, dollars
- B = Blotter cost, dollars
- T = Testing cost, dollars
- C = Total cost, dollars

Using this formula, tables were prepared showing the cost calculations used. One of them is illustrated as Table IX. Similar tables were prepared for each sizing test and each condition (Mordened and not Mordened) used. The data in the tables were plotted and are illustrated in Figure 4, which was prepared from the data shown in Table IX. The minima shown in the curves represent the optimum number of tests per sheet for lowest cost at each of the CD values shown.

A further set of plots was made from the same data, plotting evaluation cost per sample versus CD as per cent of means. Selected "high" and "low" precision values from these curves were tabulated to illustrate over-all testing costs using the three tests selected. This summary table is shown as Table X. It was evident that the use of the Morden Refiner, although causing extra working time, resulted in a lower over-all cost.

CONCLUSIONS AND DISCUSSION

In this paper, an attempt has been made to illustrate a method of determining the number of tests per unit sample tested and the number of samples necessary for any desired test precision, and a method for utilizing such data to determine lowest cost testing schedules.

Table IX

Costs - Lactic Acid Test - Not Mordened

CD Sec.	% of Mean	Tests/Sheet	a	b	aP	aM	aS	aB	bT	C
10	13.9	1	23	23	$.621	—	$17.250	$.414	$9.200	$27.49
		2	16	32	.432	—	12.000	.288	12.800	25.52
		3	13.3	40	.359	—	9.975	.239	16.000	26.57
		4	11.3	45	.305	—	8.475	.203	18.000	26.98
20	27.9	1	7.1	7.1	.192	—	5.325	.128	2.840	8.49
		2	4.5	9.0	.122	—	3.375	.081	3.600	7.18
		3	3.7	11.0	.091	—	2.775	.067	4.400	7.33
		4	3.4	13.5	.092	—	2.550	.061	5.400	8.10
40	55.7	1	3.0	3.0	.081	—	2.250	.054	1.200	3.59
		2	1.8	3.6	.049	—	1.350	.032	1.440	2.87
		3	1.3	4.0	.035	—	.975	.023	1.600	2.63
		4	1.2	4.6	.032	—	.900	.022	1.840	2.79
80	111.4	1	1.8	1.8	.049	—	1.350	.032	.720	2.15
		2	1.0	2.0	.027	—	.750	.018	.800	1.60
		3	0.8	2.3	.022	—	.600	.014	.920	1.56
		4	0.6	2.4	.016	—	.450	.011	.960	1.44

P = $.027; M = —; S = $.750; B = $.018; T = $.400

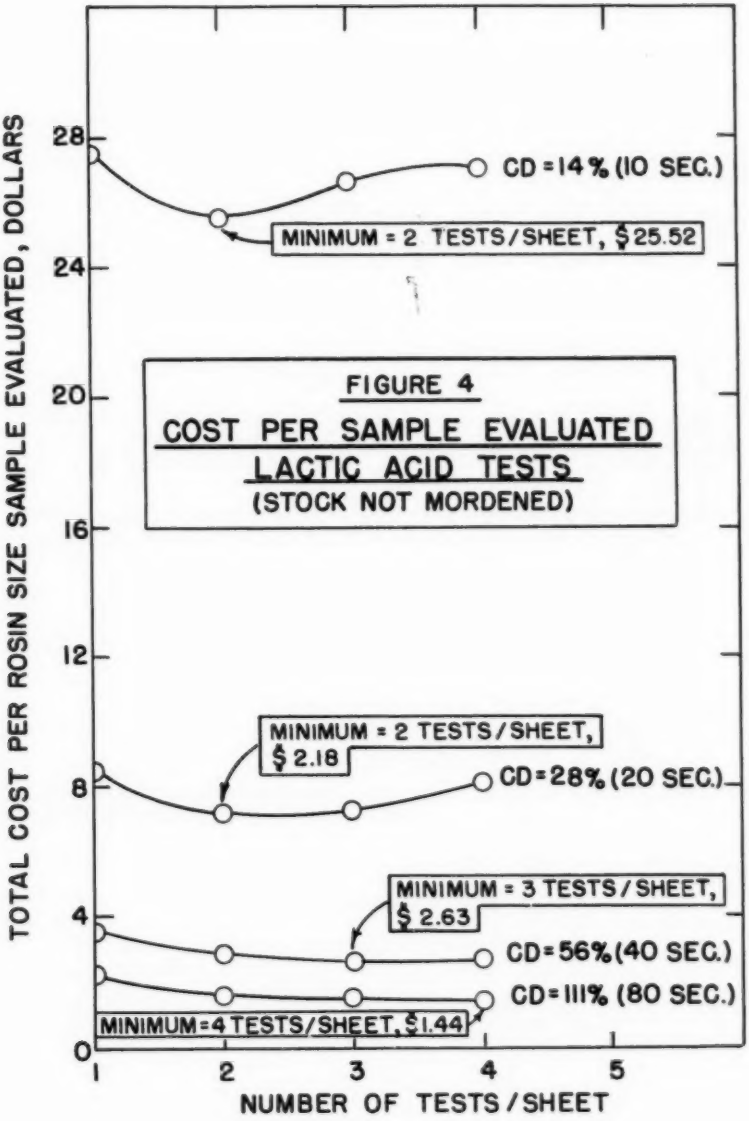


FIGURE 4

COST PER SAMPLE EVALUATED

LACTIC ACID TESTS

(STOCK NOT MORDENED)

Table X

Cost of Size Evaluations per Sample

(Summary)

Condition	Cost, Dollars							
	Ink		Lactic		Water		Total	
	10% CD	20% CD	10% CD	20% CD	2% CD	10% CD	Hi Precision	Moder. Precision
British Machine								
Mordened	$16.50	$5.20	$46.00	$11.50	$44.50	$2.40	$107.00	$19.10
Not Mordened	16.00	4.85	50.00	13.00	78.50	4.00	144.50	21.85

Although the examples used applied to paper testing, it is again emphasized that the same technique may be successfully employed with other systems involving a multiplicity of samples or preparations.

Analyses of variance were not done on the data presented herein, because conventional statistical philosophy forbids the assumption that variances of a group of data are non-homogeneous while simultaneously using them in variance analysis, in which a basic assumption of homogeneity is required. However, a number of questions arose which indicated the desirability of further work in this direction. For example, Mordening seemed to have a definite effect upon the magnitude of the errors found in the ink tests compared to those found with lactic acid data. New experiments were planned to answer these questions.

This paper represents the conclusion of this series of publications concerned with the application of statistical methods to paper sizing work.

REFERENCES

1) Wooding, Statistical Studies of Rosin Sizing Evaluation Variables: III. The Relation of Size Test Precision to Size Concentration and Two-Sidedness - Southern Pulp and Paper Manufacture, July 10, 1956

2) Wooding, Statistical Studies of Rosin Sizing Evaluation Variables: I. The Relation of Size Test Values to Size Concentration and Two-Sidedness-Pulp and Paper Magazine of Canada, October 1955.

3) Sillitto, The Number of Observations Needed in Experiments Leading to a t-Test of Significance of a Mean or the Difference of Two Means - Research Supplement II, Vol. 1, 1947-1948, Pages 520-525.

4) Wooding, Statistical Studies of Rosin Sizing Evaluation Variables: II. The Use of Variance Analysis for Rosin Sizing Test Data - TAPPI Vol. 39, No. 6, June 1956, Pages 417-425.

NOTE: References (1), (2) and (4) have been bound together and reissued as "Reprint No. 43" by the Paper Chemicals Department, American Cyanamid Company, 30 Rockefeller Plaza, New York 20, New York, and are available upon request.

ACKNOWLEDGEMENTS

For their assistance, the author wishes to thank Messrs. A. F. Blockman, C. G. Landes, R. M. DeBaun and C.A. Davis. In addition, particular thanks are due to Drs. F. Wilcoxon and C. W. Dunnett, without whose patient guidance much of our work in the application of statistics to paper technology would probably not have been done.

ABSTRACT

The cost of evaluation of a given sample of rosin size, or any other material proposed for use in a laboratory or commercial process, is dependent upon two principal factors. These are the preparation of samples and the nature of the test procedures to be used. The cost of each is related to the reproducibility of the processes involved. In most work involving physical tests, our experience has shown that the greater variation usually resides in the sample preparation steps.

If samples are in the form of a number of discrete units upon which replicate tests are to be made, as is the case with paper handsheets, the first step in determining the most economical evaluation procedure is to compute the numbers of sheets and numbers of replicate tests per sheet for selected critical differences between means of the test data. This may be done by computing "pure sheet variances" and "between sheet variances", determining the corresponding standard deviations, and using these to determine the total number of tests necessary with Sillitto's tables or by means of Student's t. For a selected critical difference, the total number of tests required may then be divided into several alternate numbers of sheets and tests per sheet. Having obtained these data, the determination of the most economical combination will depend upon the relative costs of making tests and making sheets. The optimum combination often does not correspond to the situation allowing the minimum total number of tests unless economic considerations are ignored, since the cost of testing is generally lower than that of sheetmaking.

In this paper, such calculations are illustrated, using data obtained from paper handsheets treated with rosin size and alum and produced by the British handsheet machine. The same general method may be used in any area where an analogous situation is encountered. At Cyanamid, for example, it has also been applied to the physical testing of chemically treated cellophane samples.

In a previous paper (1) the determination of the number of tests required in a given system, using Sillitto's method, was illustrated. This involved cases where only one test per sheet was used. This presentation describes an extension of the statistical procedures previously employed in that it shows the calculations necessary when more than one test per sheet on a number of sheets is used. Minimum-cost calculations may easily be made once such data are obtained, and are most useful in planning economical standard methods.

HOW TO USE INDUSTRIAL ENGINEERING METHODS STUDY FOR QUALITY CONTROL

Henry I. Matosoff
Kwikset Locks, Inc.

Each worker (production or otherwise) in performing his assigned task will either build quality characteristics into the product or will produce defective material. Which of these he has done and is presently doing will determine the quality of production. The industrial engineering approach outlines techniques that could very profitably be used to make this determination.

As the mathematicians have provided quality control with techniques of their profession, - the statistical techniques, so the industrial engineer can contribute tools of his profession - engineering analysis and methods study to achieve controlled quality. However, these techniques bring greater dividends on, a) products which are to be produced in very large quantities at a "tight" quality level, b) on products already in high volume production where quality standards have yet to be determined.

If a high degree of quality is required of a product, and we subscribe to the basic premise that "quality can only be built into a product," then a detailed engineering analysis is usually necessary in order to control the quality or production. A few definitions are in order before this thesis can be expanded:

First, quality shall be defined as the quality of conformance, that is, does production conform to blueprints, specifications, and/or classification of defects. If it does, then the product being manufactured is said to have quality built into it.

Secondly, control shall be defined as that mechanism whose primary end is the prevention of defects.

Quality control engineers will readily agree that most quality control programs are remedial, rather than preventive in nature. Because of this type of indoctrination, many quality control engineers orientate their thinking in terms of "putting out fires". Thinking of this type tends to cause quality control departments to become known as trouble shooting departments charged with applying statistical techniques to reduce high scrap or rejection rates. In fact, many authorities suggest that quality control programs should be started in troubled areas and gradually move from one such area to another until a complete quality control program is accomplished.

For the short run, "putting out fires" is efficient and causes sharp decreases in rejection and scrap rates. It also builds confidence among production people and helps them to think using statistical concepts of variation. However, for a long range, a "total quality control program" - one where quality will be repeatedly built into a product is necessary. It can be achieved by a careful analysis of the methods of production, inspection, etc. Industrial engineering analysis methodology, much the same as is used in production planning, work simplification etc. is an ideal way of making the required analysis.

If we are interested in establishing a "total Quality Control program", an analysis consisting of a detailed study of production should be determined. The analysis should indicate:

1. Where are raw material, sub-assemblies, or parts introduced into the process?
2. Where are manufacturing operations performed?
3. Where are inspection operations performed?
4. Where can defective characteristics be hidden or covered?
5. Where can defective characteristics be generated?
6. What type of gear is being used for inspection?

This paper will illustrate how such an analysis can be performed on a simple operation as soldering a connector plug.

One of the most important features in most electronic production is the manufacturing of soldered connections. It is not uncommon to find as many as 5,000 solder joints in an electronic component and a solder joint failure rate of one in 10,000 could possible result in a reliability of only 50 per cent. It is a no small wonder that some electronic manufactures strive for a rate of soldered joint failures of one in 100,000 connections. Let us suppose that this degree of quality was required. Here is how industrial engineering techniques would perform the analysis to achieve it.

THE OPERATION PROCESS CHART

First, a thorough knowledge of the existing manufacturing process is necessary and can be obtained by the use of a standard industrial engineering technique of process charting. Two types of charts are used: The Operations Process Chart and a Flow Process Chart. The Operations Process Chart furnishes an overall birds-eye description of the entire process. It shows the way component parts which are being manufactured fit together. The Flow Process Chart gives more details and shows where items are transported, where storage takes place, as well as where manufacturing and inspection operations take place.

The Operation Process Chart shows graphically where materials are introduced into the process, the order and position of operations, and the location of inspection. Two symbols are used in preparing this chart:

A large circle ○ which indicates operation; that is, steps where an object is intentionally changed, either physically or chemically or where something is added or modified to the material.

A square □ indicates where inspection is performed. Inspection is for the purpose of determing an object identification, size, quantity, function, and etc. As an example of the preparation of an Operation

Process Chart, let an electrical cable be used. Electrical cables have a relatively simple production process and contain vital soldered connections whose quality must be closely watched. The first step in making this chart is to select the component of the cable which will have the greatest number of operations performed on it. Usually it is the part with the most dimensions. The final assembly drawing or an examination of the actual part will normally indicate the proper starting point for the chart. In this case it would not matter if the wire or the connector was used as a starting point since both have few operations performed on it. However, the connector was selected.

The chart is identified by the name of the part, and its drawing number. The raw material used in making the part is then noted on the chart. Special characteristics and notations such as whether the raw material had been source inspected at the vendor should be noted also. The first step in the manufacturing on conduit covered cables is to receive and inspect the connector. A □ square is recorded on the chart. Since this is the first inspection, the figure 1. is written inside the square [1] . A brief description is then written to the right of the symbol. After this step, the connector is disassembled. A circle ◯ is placed on the chart immediately below the square to indicate this operation. This is the first operation; it is numbered (1) and described. The next operation is indicated by a (2) and again described to the right of the symbol. At this point the Operation Process Chart now looks like this:

(1) CONNECTOR

Cannon SK-C16-21 3/4
Source inspected

[1]	Receive and inspect
(1)	Disassemble connector shell
(2)	Assembly connector to holding fixture

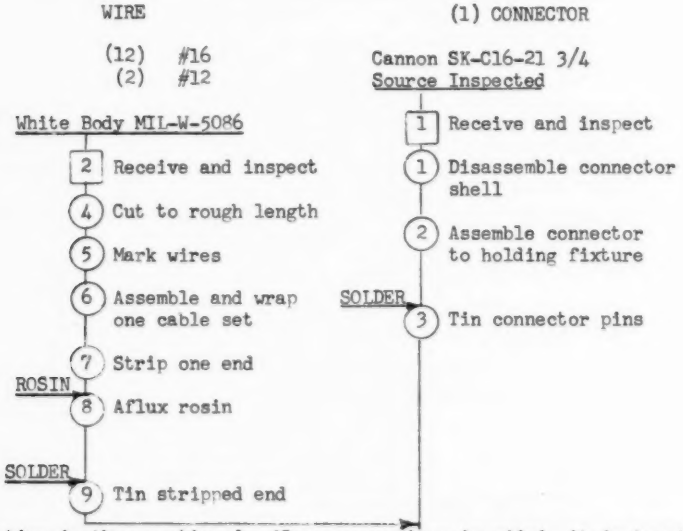
The remaining production steps in the manufacture of cables are recorded. Both operations and inspections are entered, numbered, and described. There is little difficulty in the charting of so simple of an item as a cable until we reach the point where the wire is fed into the connector. The new part (wire) must be analyzed and the steps involved in its manufacture added to the chart.

It must first be determined whether this new piece part is a simple one, or whether it is itself made up of several components. If the latter, the principal item of this piece part is selected on the same basis as was the chief component of the entire assembly. In the example,

the wire is a simple part which is purchased in reels and certified as meeting a MIL Specification. The process for manufacturing the wire is charted exactly as was the connector.

The main purpose in numbering the manufacturing steps is to identify them for later study. Therefore, the numbering of each additional component does not start at 1 but follows in sequence the last numbers used on the chart.
The chart reached operation (2) and inspection [1]. Thus, the recording of inspection and operations of the wire will start with (3) and [2]. The steps involved in the manufacture of the wire are recorded and when the wire is assembled to the connector the chart has reached this stage.

"OPERATION PROCESS CHART CONDUIT COVERED CABLE"



Charting is then continued. When a new piece is added, it is treated in the same manner as was the wire. The finished Operation Process Chart reproduced in figure 1. gives a good overall picture of the manufactured cable.

THE FLOW CHART

When the Operations Process Chart is completed, the Flow Process Chart can be started. This chart will analyze and focus its attention on one part or component of the manufacturing process and examine it is detail. A Flow Process Chart is more detailed than an Operational Process Chart. The Flow Process Chart used for Quality Control is primarily established to help determine where to locate inspection stations and what characteristic to inspect. The Flow Process Chart will show

the order of all operations, inspections, transportations and storages. It will also indicate where dimensions or characteristics are generated and where they are hidden or covered. The Flow Process Chart is prepared with the aid of the Operation Process Chart, the component drawings, and the item in question (cables).

The chart is headed and identified with the part being studies, its drawing number, and description of the first and last manufacturing step described by the Flow Process Chart. Symbols are noted on the lefthand side of the Flow Process Chart. Two of these, the large ◯ circle and the square □ were used in the preparation of the Operation Process Chart and have the same meaning here. The small circle ○ stands for a transportation, the moving of an object from one place to another. The triangle △ represents storage, a stage in the production process where the material or object is intended to remain at one place.

In addition to the symbols, the chart form also contains space for the description of each step and space for the answering of two questions concerning each step. Figure 2. shows a Flow Process Chart for the wire which is being used in the cable assembly.

One of the first steps in the production of a cable is the preparation of wire. The wire is stored in racks and therefore the triangle is blackened ▲ and described appropriately. Since the two questions on the Flow Process Chart are not pertinent at this step, they are left unanswered.

The next step is a transportation. The small circle ● is blackened and joined to the triangle which is on the line above. The step is then described and examined. At the next step there is an operation and 4 is entered in the large circle (4) identifying the operation as corresponds to that marked (4) on the Operation Process Chart. Each of the operation steps must be analyzed with particular care since characteristics or dimensions are usually generated at these steps. The first question on the form asks which characteristics are generated at this operation. At a later stage the wire is cut to its final length. "Is this measurement the final length?" If it was, then the particular characteristic of wire length is generated and the description of the characteristic would have been entered on the Flow Process Chart. However, in this example, the cut is a rough one and later in production the final wire length is made, thus, a characteristic is not generated in (4).

In answering the question concerning where characteristics are generated, the entire process must be taken into consideration. The Quality Control Engineer should use not only his own knowledge of production, but also the knowledge of the operator or the foreman. The key to the proper answer to the question where are characteristics generated is, "Can final inspection of the characteristic be performed at

this point?" If it can, the characteristic is said to be generated.

The next step in the production of the wire is the transportation to the marking machine which is entered on the chart as a transportation. Next step, is an operation, marking of wire, and requires close study. The Quality Control Engineer must satisfy himself by answering such questions as "What happens here"? What operations does the machine perform? How does each operation affect characteristics on the item being manufactured"? The answers can frequently be obtained by watching the manufacturing process and the worker doing the work. For example, the difference between a rough cut and a finished cut is usually apparent. In some cases it may be necessary to question the worker or his foreman in order to obtain the answers from experienced men.

As the Flow Process Chart is filled, attention must be paid to the question listed on the form. The first of these, which concerns the generation of characteristics, has already been discussed. The second question asks whether a characteristic is hidden or covered. The following definitions are applicable:

Covered: A characteristic that could only be inspected by a later non-destructive disassembly operation. Example, thread characteristics of an assembled nut and bolt.

Hidden: A characteristic that cannot be inspected by a later disassembly operation, because the disassembly operation would destroy the parts involved. Example, the finish of a surface which later is painted; bevel of an edge which later is welded.

When the question of whether a characteristic is hidden or covered is answered "yes" the characteristic in question should be identified. An example of a hidden characteristic occurs at operation (9) in the manufacture of a cable. This operation of tinning may cover wires which may have been knicked in the stripping operation or which may have been stripped short.

The Flow Process Chart is completed by charting the steps in the manufacturing process to the point indicated by "Chart Ends". A separate Flow Process Chart should be prepared for each component. In this way a series of charts will be developed covering a manufacture of the items under study. The completed charts must be carefully checked. It is most important to make sure that the generation of every characteristic has been entered in the proper step. When the Flow Process Chart has been completed and checked, the basic data concerning the manufacturing process has been collected and further analysis as to the location of inspection stations may begin.

"THE LOCATION OF INSPECTION STATIONS"

After the actual manufacturing processes are studied, through the use of the Flow Process Chart and the Operation Process Chart, the Quality Control Engineer is ready to determine where to locate points of inspection and what characteristics should be inspected at these inspection points.

At this point he should also be in a position to suggest changes

in manufacturing or inspection methods. In the example being used, the new inspection station location accomplished all the methods changes desired.

Before inspection stations can be located, several management decisions are necessary. First management must decide whether the "Inspector" or the "production" approach of inspection shall take place. The "inspector" approach is that an item be inspected as late in the manufacturing process as possible. This is the type used by the Military for purposes of acceptance inspection. The "production" approach usually is that the inspection should take place as soon after an operation as is possible. Normally the Quality Control Engineer will advise management on the use of a modified approach, that is, to inspect as late in the process as is possible, but before a vital characteristic is covered or hidden. The cost of having discrepant material in the manufacturing system will dictate management's decision. In any case, the Operation Process and Flow Process Chart will help crystalize the area of decision.

Secondly, management must consider the cost of inspection. This will be decided by the complexity of the item and the quality which is desired by the customer. The AQL, AOQL, and/or the Classification of Defects can greatly assist management in this decision. In the case of the solder connection, the cost of inspection was dictated by the AOQL of one defect in one hundred thousand connections.

Thirdly, management must decide the rate of its production. If the rate of production is ten units, different methods, techniques, or controls will be used then if the rate were ten thousand units.

Let us suppose that in the example of the solder connection management's decisions is reliability at all other cost then quality dictates all other considerations.

If this is the case a modified inspector's approach is used. (Inspect as late as possible, but before characteristics are covered or hidden). Using this approach, normally both the desired quality and a reasonable cost of inspection would result. If it didn't, the production approach (that is, to inspect as soon after an operation as possible) would have been instated.

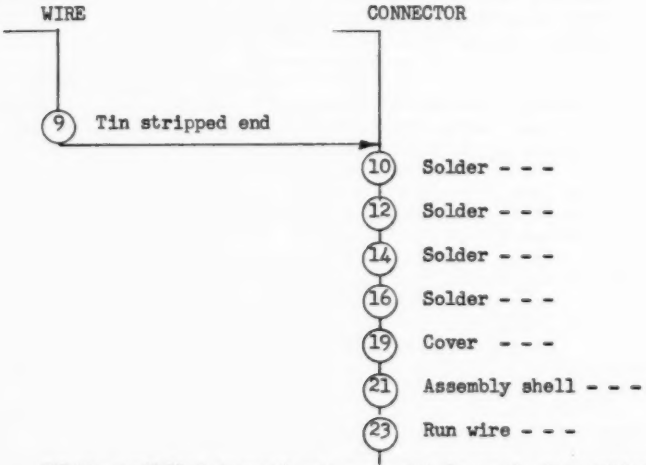
WHAT CHARACTERISTICS MUST BE INSPECTED

1. Improper wire size
2. Improper wire length (short or long)
3. Continuity of wire reels
4. Wire identification number improperly spaced
5. Bulges or tears in wire insulation
6. Wire stripped too far
7. Wire stripped too close
8. Wire not stripped square
9. Outside diameter of wire reduced
10. Presence of flux or rosin in solder
11. Cold solder
12. Excessive solder
13. No solder
14. Too little solder
15. Burned insulation
16. Missing wires
17. Duplicated wires
18. Wrong plug
19. Damaged plug
20. Crushed pins
21. Cracked plugs or inserts
22. Wire not flush in pin
23. Reversed wires
24. Presence of foreign material
25. Wrong size of vinyl cover

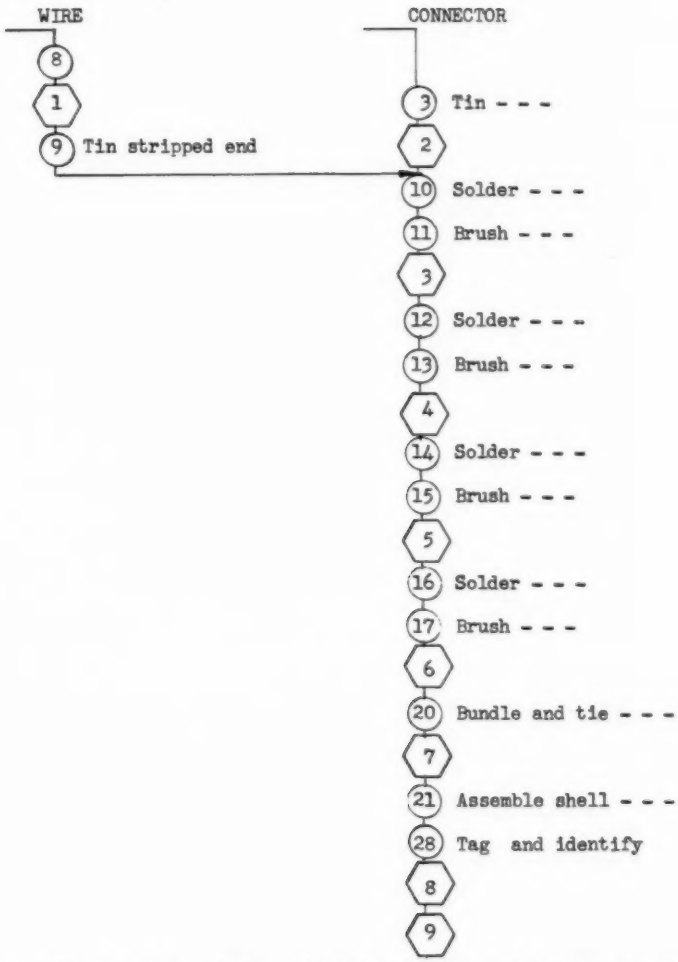
26. Loosely tied bundle
27. Improper length of vinyl cover
28. Missing, loose, or damaged plug screws or plug threads
29. Improper cable length, or size
30. Improper cable identification
31. Damaged or crushed conduit
32. Improper location of plug insert, relative to plug
33. Ferrule Improperly assembled to cable
34. Improper lug size
35. Less than - - - ohm on megger
36. More than - - - ohm on continuity test
37. Lugs not fastened securely
38. Wire on lugs extended too far
39. Identification not clear

WHERE TO INSPECT

Finally the Quality Control Engineer must determine where in the manufacturing process these potential defects should be inspected. This is done by analyzing question 2 of the Flow Process Chart. "Do characteristics become hidden or covered?" In the example of the cable, this question was answered postively in the following places:



If the modified inspector's approach is used, inspection must take place before characteristics can be hidden or covered, since potential defects may have been generated. The ⬡ indicates where our inspection stations, using modified inspector's approach, were located.



Since you cannot inspect that which has not been produced, only the characteristics which are generated before they reach the inspection station can be included on the inspection list. The following is a list of characteristics that must be inspected at each inspection station.

INSPECTION (1)

Improper wire size	ref. W.O., wire gage or compare visually
Bulges or tears in insulation	visual
Stripped ends more than 1/4" or less than 3/16"	rule
Wire knicks at stripped end	visual
Stripped ends not cut square	visual
Improper wire length	ref. drawings, rule

INSPECTION (2)

Stripped ends cold soldered	visual
Stripped ends excessive soldered	visual
Stripped ends not soldered	visual
Burned insulation	visual
Excess of rosin or presence of foreign matter at stripped ends	visual

INSPECTION (3) (4) (5) (6)

Wrong wire in pin	visual
Wires missing	visual
Duplicate numbered wire	visual
Crushed plug pins	visual
Pin filled with solder (female connection)	visual
Excess solder	visual
Too little solder	visual
Cold solder	visual
Dirt, chips, excess of flux or foreign matter present	visual
Cracked plug or inserts	visual

INSPECTION (7)

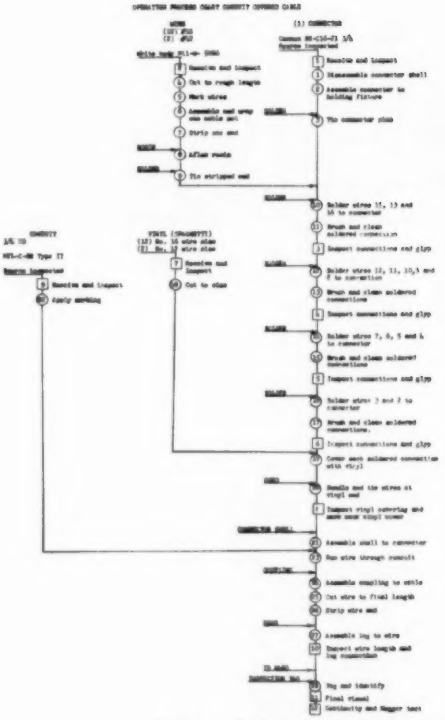
Improper vinyl size, fits loosely over wire	feel
Improper vinyl length	visual-rule
Missing vinyl covers	visual
Cable wires not firmly bound	feel-visual

INSPECTION (8)

Wrong plug 1 & 2 used	ref. parts list visual
Missing coupling parts on plug 1 & 2	visual
Improper cable size or length	ref. drawing rule
Improperly assembled cable and plug 1	ref. drawing visual
Improperly marked cables	visual
Cable bands missing	visual
Crushed or damaged cable	visual
Ferrule improperly assembled to cable plug 1	ref. drawing
Ferrule damaged on plug 1 & 2	visual
Improper location of insert on plug ref. drawing	rule 1/16 graduation
Loose, missing, short, burred or stripped threads on plug 1	visual
Excess solder	visual
Too little solder	visual
Cold solder	visual
Pin filled with solder	visual
Dirt, chips, excess of flux or foreign matter present	visual
Wrong wire in pin (plug 2)	visual
Wire missing (plug 2)	visual
Duplicated numbered wire	visual
Crushed pins on plug 2	visual

INSPECTION (9)

Cable plug does not fit board plug	visual
Megger reading less than - - - megohms	visual
Continuity test more than - - - ohms	



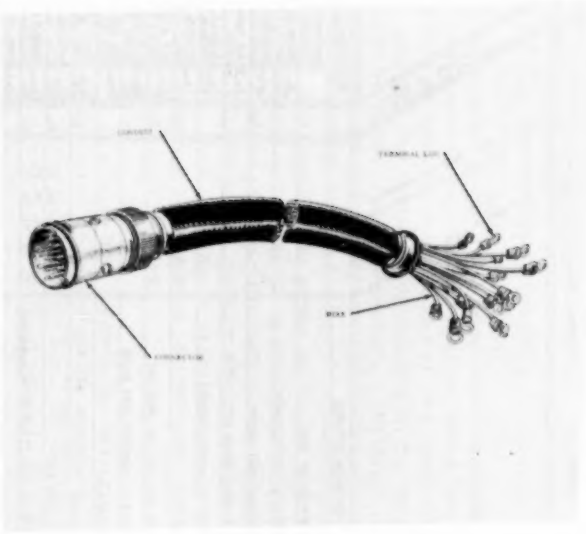


Figure 1 Operation Process Chart Conduit Covered Cable

FLOW PROCESS CHART

PART OR DRAWING Nº. Wire

CHART STARTS "Storage area for wire"

CHART ENDS "Assemble Shell

PAGE OF P85

DATE

BY H. I. M.

OPERATION
TRANSPORTATION
CONTRACTORS INSPECTION
STORAGE

DESCRIPTION	What characteristics are observed at this step	Is it inspected 100% hidden or observed	
Storage area for wire			
Transport to cutting table			
Cut wire to rough length	1, 3	No	
Store in wire trough			
Transport to marking machine			
Mark wire	1, 3, 4, 5, 6	No	
Assemble and wrap one cable set		No	
Transport to operators work bench			
Strip one end			
Aflux rosin		No	
Tin stripped end	11, 12, 13, 14, 15	Yes	
Solder wires 15, 13 and 16 to connector	16, 17, 18, 19, 20, 21, 22, 23, 24, 29	Yes	
Brush and clean soldered connections		No	
Inspect connections and glyp			
Solder wires 12, 11, 10 9 and 8 to connector	15, 17, 18, 19, 20, 21, 22, 23, 24, 29	Yes	
Brush and clean soldered connections			
Inspect connections and glyp			
Solder wires 7, 6, 5 and 4 to connector	16, 17, 18, 19, 20, 21, 22, 23, 24, 29	Yes	
Brush and clean soldered connections			
Inspect connections and glyp			
Solder wires 3 and 2 to connector	16, 17, 18, 19, 20, 21, 22, 23, 24, 29	Yes	
Brush and clean soldered connector			
Cover each wire with vinyl	25, 26, 27, 28, 29	Yes	
Bundle and tie wires at vinyl end			
Inspect vinyl covering			
Assemble shell to connector			

POSSIBLE DEFECTS

1. Improper wire size
2. Improper wire length
3. Holes or tears in insulation
4. Identification number improperly spaced
5. Wrong wire marked
6. Wrong quantity of wire marked.
7. Stripped too far
8. Stripped too close
9. Stripped ends not square
10. Wire knicks
11. Cold solder
12. Excessive solder
13. No solder
14. Burned insulation
15. Wire not flush into plug pin
16. Pin filled with solder
17. Wrong plug used
18. Reversed wires
19. Wires missing
20. Cracked plugs or inserts
21. Duplicated wires
22. Wrong plug
23. Loose, missing, short burred or stripped screws
24. Improper vinyl size
25. Improper vinyl length
26. Missing vinyl covers
27. Cable wires not firmly bound
28. Dirt, chips, excess of flux or foreign matter

Figure 2 Flow Process Chart

FINDING THE FACTS BEFORE SETTING THE SPECIFICATIONS
II METHODS

Julian Harrison Toulouse
Owens-Illinois Glass Co.

Let us first establish one or two fundamental premises, first - that almost every business organization is at the same time both a buyer and a seller of some form of product or raw material. My company, for example has thousands of customers, but we ourselves are customers to hundreds of other suppliers. The same is true of every organization here represented.

We follow specifications for the things we buy as well as the things we sell. Because of manufacturing information required, the blueprint details of a bottle for the most meticulous customer differ little from those of a bottle for a customer who may not know that blueprints are needed. And the same is probably true in any field of manufacture, from mouse traps to diesel locomotives, from lipstick to the mechanisms for nuclear fission.

Secondly, each of us sits on both sides of the bartering table. We must listen to the same details about the quality standards of the things we sell, which we use about the things we buy. We desire high performance of the things we buy - we must agree to give high performance in the things we offer for sale. The inter-relation of the buying and selling aspect of each company becomes like a piece from a jig saw puzzle, or like the Yang and Yin of Chinese philosophy, each part finding its harmonious counterpart.

Then too, let us not think of specifications as being part of only a producer-consumer relationship. Many more specifications exist entirely within an organization, a matter of greatest concern only between departments of the company. Take an automobile, for example. Hundreds upon hundreds of specifications are necessary for the component parts of their assembly, and these specifications (except to the suppliers of materials) largely never get outside the place of automobile manufacture. What passes for "specifications" in the sales literature are merely recitations of a few standard dimensions, with a hint as to performance.

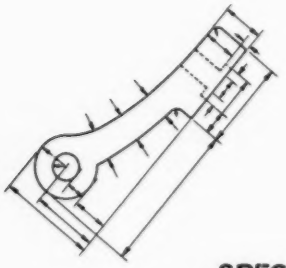
Whether hinted or spelled out in detail, performance is exactly what is exchanged over the trading table. It is fundamentally true that the only excuse for specification in any regard is to outline some necessary phase of performance. Herein lies the reason, the basis, and the whole excuse for any expressions of specification, of tolerance, or of descriptive intent. All too many specifications go beyond this point into the field of conjecture, coercion, or fantasy to the end that difficulty is made certain, costs pyramid, waste is inevitable, and the specification itself becomes an anathema.

To this problem, I once made the plea - "find the facts before setting the specification", in a speech by that title (1). This plea is now renewed, and my part of today's program is to tell by example some of the methods by which facts can be found. Since I so completely believe in the use of the statistics of probability as a tool for study, the methods I will mention will all be statistical. In a sense I am using the need for facts to show how statistical approaches can be made.

I want to make clear at the outset that I believe too few specifications arise out of the facts of performance, which should be the only reason for a specification. Too many specifications are based on concensus and compromise of opinion, or on mutual distrust, or out of the game of claim and counterclaim. I believe that specifications should be arrived at by free discussion of the needs, and when so established, should be met and enforced.

Dr. Gailliard has detailed the correct idea of specifications - let me add some incorrect methods of arriving at a specification, by name and description:

Fig. 1. The Fine Print Method
Too many specifications include many details not necessary for actual use of the article being specified. Blue prints for manufacture may have many dimensions necessary to develop the article. Too often a specification for compliance includes every dimension on the blue print. Don't confuse a specification detail with a working drawing.



FINE PRINT

METHOD

SPECIFYING DETAILS

NOT NEEDED FOR PERFORMANCE

BUT DEMANDING COMPLIANCE

Fig. 1

Fig. 2. The Umbrella Method
When specifications are the result of a group discussion of standards, they are often drawn so that "nobody is left out". When concerned with a group of producers, it is often set so as to include the weakest number of the group. When concerned with a group of consumers, it is often set to include the tightest specifications desired by any one member.

UMBRELLA

METHOD

DESIGNED SO THAT EVERYONE

CAN GET UNDER

Fig. 2

SHOTGUN WEDDING

METHOD

A FORCED UNION ARISING OUT OF A

DOMINANCE BY EITHER SIDE

Fig. 3

Fig. 3. The Shot Gun Wedding Method
Dominance of one side of the agreement results in a forced union. Compare the terms "buyers market" and "sellers market". Each produces a different idea as to specifications.

Fig. 4. The Keeping Up With the Jones' Method
Sometimes specifications are copied from another source without regard to the suitability.

KEEPING

UP WITH THE

JONES' METHOD

COPYING ANOTHER, AND PERHAPS

UNRELATED, SET OF SPECIFICATIONS

Fig. 4

Fig. 5. The Go You One Better Method
Setting tighter specifications, just to gain or maintain a reputation for toughness.

GO YOU

ONE BETTER

METHOD

ELABORATION AND EXACTINGNESS

FOR PERSONAL GRATIFICATION

Fig. 5

Fig. 6. The Fiction Writer Method
So far removed from the needs as to have no resemblance to "persons or events".

FICTION WRITER METHOD

ANY RESEMBLANCE TO PERSONS OR EVENTS IS ENTIRELY ACCIDENTAL

Fig. 6

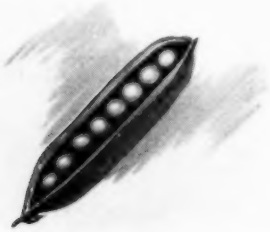
Fig. 7. The Ostrich Method
Entered into by burying one's head in the sands unaware of what is going on, completely ignoring the subject.

OSTRICH METHOD

BURYING ONE'S HEAD, UNAWARE OF WHAT IS GOING ON

Fig. 7

Fig. 8. The Peas in a Pod Method
Desired by the perfectionist. He cannot believe in variation - all articles must be exactly alike (but did you ever look at a pod of peas?). This man sets tolerances as the smallest number he can think of, divided by two



PEAS IN A POD METHOD

BELIEVING THAT THINGS DO NOT VARY EXCEPT BY CARELESSNESS

Fig. 8

Fig. 9. The You Go Your Way Method
Complete disregard of specifications. Often set tight by engineering, hoping that production will come close, or knowing that production will set its own standards.

YOU GO YOUR WAY METHOD

ENGINEERING SETS TOLERANCES TIGHT BECAUSE PRODUCTION WILL MODIFY THEM IF IT SEEMS NECESSARY

Fig. 9

Fig. 10. The Hairline Method
Tolerating no deviation whatever from tolerance even by the most minute amount. No concept of "Average Out-going Quality" and statistics. No "Material Review Committee" to decide on borderline quality.

HAIRLINE

METHOD

THE AX FALLS ON A HAIRLINE DEVIATION

NO MATERIAL REVIEW COMMITTEE

Fig. 10

Fig. 11. The Accent the Positive Method
Sometimes based on what Irvin Bross in his "Design for Decision" calls a "selective amnesia for the facts" - largely exaggeration and cover up by both sides - emphasizing only that which is favorable to the side concerned.

ACCENT THE

POSITIVE

METHOD

A GAME OF EXAGGERATION AND COVER-UP.

"SELECTIVE AMNESIA FOR THE FACTS"

BROSS

Fig. 11

Fig. 12. The Damn the Torpedos Method
No regard for the dangerous waters of the increased procurement cost to an unnecessarily tight specification, or the line production cost of using materials procured under loose or inadequate specification - can go either way, in spite of an intermediate channel.

DAMN THE

TORPEDOS

METHOD

SETTING SPECIFICATIONS REGARDLESS OF MATERIAL OR PRODUCTION COSTS

Fig. 12

If accurate costs are obtained we can make a mathematical comparison to find the best-cost-specification. Line costs decrease with less variable material, and increase as the variability increases, as shown by the curve labeled "Cost of use increases with loose specifications". Procurement costs go up with tighter specifications, simply because it can cost more to make such material, as shown by the curve marked, "Purchase cost increases with close tolerances".

The real cost is their sum. In this chart they are added graphically to give the curve labeled, "The combined costs pass through a minimum point which can be calculated". This minimum is the best specification for the user, and it is shown that it can be missed slightly without serious difficulty. If each curve could be set up mathematically with "X" as the variation in tolerance the minimum point can often be calculated by the Calculus. The graphic method differs only in the ability to make determination at enough points to determine a smooth curve, whereas the Calculus smoothes out irregularities.

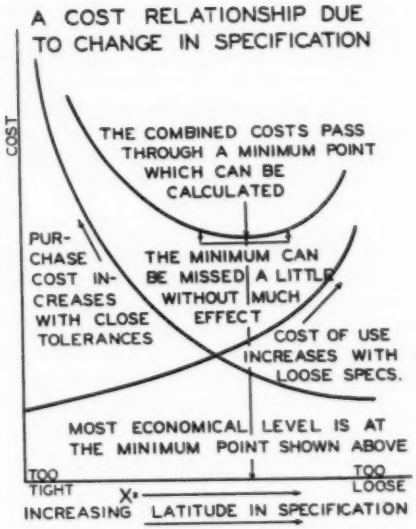

A COST RELATIONSHIP DUE TO CHANGE IN SPECIFICATION
COST
THE COMBINED COSTS PASS THROUGH A MINIMUM POINT WHICH CAN BE CALCULATED
PUR-CHASE COST IN-CREASES WITH CLOSE TOLERANCES
THE MINIMUM CAN BE MISSED A LITTLE WITHOUT MUCH EFFECT
COST OF USE INCREASES WITH LOOSE SPECS.
MOST ECONOMICAL LEVEL IS AT THE MINIMUM POINT SHOWN ABOVE
TOO TIGHT
TOO LOOSE
INCREASING
LATITUDE IN SPECIFICATION

Fig. 12-A

Now in detailing some typical statistical approaches toward finding the facts before setting the specification, I can only use a few illustrations. Each business has its own problems, to which these can only be adapted by one knowing the industry. I will first mention a few covering the product, then the consuming side (which may be a customer or another department in the same business organization), and finally the producing side (even if only the preceding company department). The methods used are of course not the exclusive prerogative of either producer or consumer, nor are the things that are measured.

The Product

1. Distribution of Measurements of a detail in relation to a specification.

Let us first illustrate what is called "frequency distribution", for the benefit of the non-statistical. Suppose we took one hundred dimes and measured their thickness. Suppose we set up a pile for each increment of thickness, letting the number of dimes in each pile therefore illustrate the relative "frequency" of each thickness. Such an array is called a distribution, whether an actual assembly of the parts, or by a graph in which the heights of bars or lines, or the positions of a series of dots exemplify the situation.

Such a "frequency distribution" so often presents a regular bell-shaped contour that the individual piles or bars can be shown as though connected by a smooth curve - the "normal curve of distribution" - often, as I will do in the rest of this presentation, all this is shown simply by drawing the curve itself, carrying the implications that measurements have been made, and the curve merely connects, graphically, the tops of the piles, with its high point representing the approximate average, and its tails fixing the extent of variation from that average.

The graph of the thickness measurements of those 100 dimes, or any given number of measurements of any other thing, shows visually two of the fundamentals of all probability statistics - a measure of the average, usually symbolized by $\bar{X}$, and a measure of the spread of the curve (sigma), wherein a distance of three times sigma on either side of the average usually includes 99% or more of all of the measurements. This useful measurement of variability is perhaps the most important mathematical tool of the statistics of probability.

THICKNESS DISTRIBUTION OF 100 DIMES

Fig. 13

2. Relation to Production and Specification Tolerances

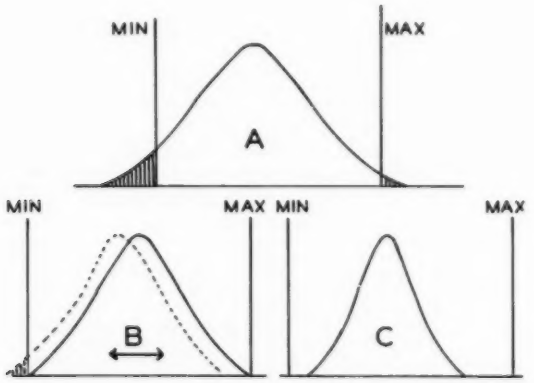
Now let's symbolize a relationship between a product of a manufacturing line, and the specification tolerance limits by using the normal curve and appropriate limit lines. The fact that the curves are broad or narrow does not matter - only their approximate shape.

Figure 14-A shows a curve extending beyond the tolerance limits on both sides. It indicates great variation with respect to the tolerance limits. This is an untenable situation. Either the fringe lying outside each tolerance must be removed by measuring in detail every item produced (a costly procedure at best), or a new and better production method must be found (often a matter of research and delay), or some way must be found to make wider limits feasible by altering the method of using the product.

Figure 14-B shows the curve just clearing inside the limits on either side. First thoughts might be that this is perfect, but it is far from ideal. No operating tolerance of any kind is possible, and the dotted lines show how the product can be partially out of specification by only a slight shift of the average.

Here is where many specification writers drop into a pit-fall. They measure a few items, which cannot contain both extremes of the product, or even one extreme, and then set limits about the range of these few samples. For example, the probable variation of the product from which a sample of 5 is taken would be on the order of two and one-half times the spread (or range) in measurement of the five items. The average spread of five items randomly sampled is shown by the short, double arrow.

Figure 14-C shows a more comfortable situation, one which allows some manufacturing leeway, some degree of tool wear or similar variation, and some degree of latitude for combining the product of two or more machines or spindles. Under this condition one can expect good compliance with a specification.



SOME DISTRIBUTIONAL ASPECTS OF SPECIFICATION

Fig. 14

3. An area study

Other measures of product condition can be used. Chart 15-A shows a common condition of paper board, as shown by Mullen Tests, six inches apart in every direction, first as originally taken, then as combined into averages of each group of 9 which forms a square. The average of each group of 9 is placed in its center on Chart 15-B with cross-hatching to show zones. A different idea as to the board is obtained from the area averages. Possibly this explains, and gives an idea as to, the decision "when is paper wild?". It also shows why a true random sample is needed for any evaluation.

Grouping the Mullen test samples in one area would give erroneous results, yet so often this is done. A completely different evaluation would be made if six adjacent samplings were made in either the low or the high test areas.

Thus we have given three different kinds of study of a product isolated from production - a frequency distribution, a specification adherence probability, and an area study. Each might give an answer to the question of adherence to specification, but they are only a few of many examples of possible product variation studies.

MULLEN VALUES: INDIVIDUAL
ONE TEST IN EACH SIX INCH SQUARE

245	240	227	253	260	223	265	246	230	241	225	190
220	228	244	240	233	215	248	255	225	185	205	200
215	240	233	239	220	232	252	257	240	240	245	220
267	235	200	230	260	200	211	233	240	205	233	194
240	220	213	250	205	255	233	238	228	220	190	202
230	209	228	217	240	212	275	229	194	232	219	215
235	208	245	233	239	225	189	220	210	190	183	200
245	230	220	235	246	226	228	193	208	240	200	215
216	255	225	200	233	200	260	239	235	235	205	205
255	260	251	237	265	248	220	228	260	220	215	190
250	245	219	237	260	236	235	250	231	224	215	252
245	230	260	240	202	255	220	220	233	212	221	220
260	253	215	232	239	239	242	200	253	201	218	192
238	233	233	255	240	245	233	248	258	227	200	215
232	228	235	252	263	253	244	260	217	212	210	205

Fig. 15-A

MULLEN VALUES: AVERAGES OF NINE
FROM ONE TEST IN EACH SIX INCH SQUARE

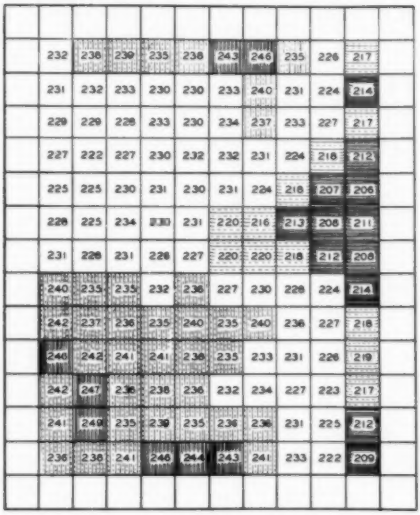


Fig. 15-B

The Consumer

1. Line Operation

We can only hint at what might be studied in a line operation. Perhaps you have already made up your mind as to the cause, and study only that one cause over a short run. If you do, you may be badly out of line; you may be gathering only those data which prove your point - one form of the selective amnesia already mentioned.

Set up an extended study recording all the non-conforming units found, separate them as to classes, and tabulate the stage in the process, forming a table similar to Fig. 16.

Table 1

A process line survey, over an extended period

Class	Stages in the process 1	2	3	4	Sum	% of Grand Total
	Non-conforming units					
1	606	1093	235	319	2018	45.23
2	227	208	84	99	534	11.97
3	38	46	5	5	89	1.99
4	38	37	7	28	103	2.31
5	27	27	25	67	121	2.71
6	84	64	69	326	474	10.62
7	74	97	35	236	407	9.12
8	26	218	74	204	448	10.04
9	3	2	9	107	112	2.51
10	0	7	0	1	8	.18
11	0	0	0	3	3	.07
12	14	33	54	98	145	3.25
					Grand Total	
Sum	1137	1832	597	1493	4462	-
%	25.5	41.1	13.4	33.5	-	-

Now for a bit of statistical slight of hand. Class 1 appears at all four stages in the process, but mostly in stage 2. Is the reason due to raw material or to something in the stage? If the former, perhaps there is something to do in the way of a specification study; if the latter, it would be better to get to work on the properties of the stage, the machinery, the process or the like, in addition to the specification.

The statistician here applies a significance test - a mathematical procedure which, within limits, answers the question - are these two sets of data really different, or could they reasonably come as two separate samples from the same system of causes? In this case it can answer the question: Is the number of non-conforming units different enough that the controlling cause is the stage, or is there enough similarity that the cause is something to do with the class, or something common to all items of that class as raw material, etc.

There are many formulas which can be used to estimate the significance of a pair of results. Since this production was of a large number of units and both figures to be tried come out of it, we might use the very simple formula for numbers in an area, although it is not the best to apply.

$$t = \frac{C_1 - C_2}{\sqrt{C_1 + C_2}}$$

Since Class 1 gave 606, 1093, 235, and 319 as the figures, we can choose any two. A good start is to take the largest and smallest.

$$t = \frac{1093 - 235}{\sqrt{1328}} = \frac{858}{36.44} = 23.54$$

Now, since it can be statistically shown that any value of t greater than 2.6 is an indication that the same system of causes could not reasonably give both of the values concerned, we can concentrate on stage 2, and omit stage 3 for the moment. Let's now see if stage 1 is significantly different than stage 2.

$$t = \frac{1093 - 606}{\sqrt{1328}} = \frac{487}{36.44} = 13.36$$

Thus we have found that the data for stage 2 is worth looking into. We still do not know if it is the raw material, the stage 2 process, or a combination of several causes but a planned experiment could find the answer.

The first observation you may make is that all this was quite obvious since stage 2 in class 1 contained one-fourth of all the non-conforming production. True - but seldom do we find such data available. This table was developed when the producer went into the consumer's plant and got it. The consumer hadn't even thought of such necessity for good data! This is the real point of the discussion. A study pin-pointed the reason, and the reason in this case turned out to be machinery maintenance - not a need for a specification change!

A further step, and better, is to divide the data collecting period into sub-groups. The preceding was a three month study. Let's look at a comparison of three weeks in another study.

While we could use tests of significance, the above can be judged by another technique - analysis of variance. Without going into details, it showed that there was no real significance between the days of operation, but that there was real significance in the classes of non-conformity. Also, stage 1 was not much changed day-to-day; any improvement here would be a matter of material or process and could be permanent. Stages 2 and 3 were variable, meaning that some cause was operating which changed from day to day. It illustrates the possibility of what the statistician would call an

assignable cause. In this case it had much to do with the ability of inspectors.

Table 2

A three week process line study

Stage	Days 8/4-8	8/11-15	8/18-22	Total
	Non-conforming units			
1	188	288	208	624
2	15	79	49	143
3	167	151	119	437
Total	370	458	376	1204
Units Processed	473610	670700	608310	1752620
	Converted to percent by weeks			
				Avg.
1	.0397	.0429	.0342	.0356
2	.0032	.0118	.0081	.0082
3	.0353	.0225	.0196	.0249
Total	.0782	.0772	.0619	.0697

We gained a lot of ground quickly in respect to the process, but it was done by a planned experiment involving separation of causes, and not by overall data haphazardly arrived at. Analysis of variance is only one of several "designs for experiment" which give more credible facts for decision purposes. They are what statisticians call "Decision Functions" because they measure the probability that an assumed answer is the correct answer. And what business man would not give his right arm to know how to be correct in nearly every decision!

A final process study is shown as Table 3, where five similar plants used the same methods and supplies.

These data are in percent of the total non-conforming material found in any one plant and merely illustrate this point: That the difference between plants was more often the cause of non-conformity than the raw materials.

Table 3

Five plants using the same materials

Plant	Stages				
	1	2	3	4	5
	% of all the non-conforming units involved				
A	62.2	4.4	6.5	25.2	1.7
B	40.0	11.6	26.3	22.1	0
C	45.3	9.1	18.6	22.4	4.6
D	56.1	14.2	9.1	17.8	2.8
E	17.8	17.8	32.2	17.8	14.4
Avg.	52.7	11.1	12.7	20.5	3.0

2. A performance test

Now we can come to a performance test, typical of many, but in which a new principal of design in corrugated containers is possible. This is a study reported in the Journal of the American Society for Testing Materials, and involves curvilinear correlation (2).

The logic of the study was this: One of the tests of a corrugated package is the Conbur test, or simulated damage in a freight car by a sudden change in speed, usually deceleration as in switching and bumping another car at slow speeds. It has reached a point of high confidence as a performance test. If damage occurs, it must be through the crushing of the corrugations to allow an almost uncushioned impact. Figure 16 illustrates this.

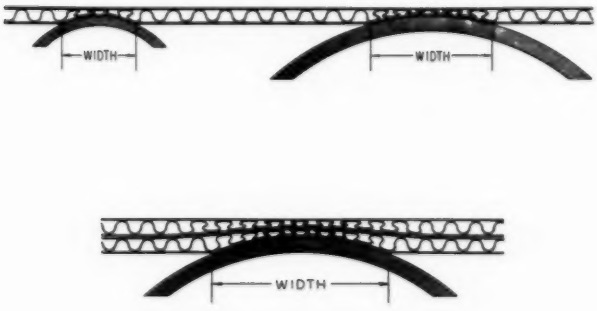


Fig. 16

One measure of corrugated board quality is the force necessary to crush the corrugations. Perhaps we could use the crush test in a correlation with the conbur test, thus relating a laboratory test, for specification purposes, to a performance test. We could then specify the strength of the board necessary to carry a product safely.

We had data as to the crush test of a number of boards and as to the conbur test of a number of products packaged in them. Figure 16 shows what happens when a bottle crushes half way into the corrugations - an area is contacted. We are dealing in this illustration with non-test "A" flute.

The formula for a hyperbola seemed to be applicable because it would be compatable with no-crushing and therefore no damage when the load comprised of the weight of bottle and content approached zero, and no resistance when the load approached infinity. Therefore, we equated the load concentration of the article contained (its weight divided by the area in contact with a half-crushed corrugated wall) with the conbur foot-fall result, using the method of least squares, and the formula, Log X = n Log Y - Log K, and obtained the curve for a specific material shown in Chart 17.

Now we are ready for design matters. Suppose it were decided that a minimum of 20 foot-falls were to be desired. The half-gallon and gallon bottles could not be packed in this non-test medium. The others could be packed. Thus one could specify this board, among

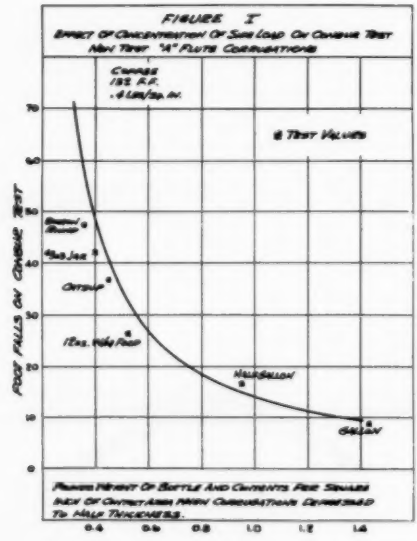


Fig. 17

others, for use, limiting the size of the container with which it could be used. A stronger board, probably more costly, would be specified for the gallon; it would cost too much to use where it was not needed.

This, then, illustrates a use of curvilinear correlation to find the facts before setting the specification.

3. Additive Tolerances

One concern of specification writers is the overall tolerance of an assembly of several parts. Thus, if several piece-parts are assembled, it is a popular idea that the tolerance of the assembly becomes the sum of the tolerances of the individuals. In other words, people assume that if 5 items are assembled together and each has a tolerance of .002", then the resulting variation of the assembly would be 5 times .002 or .010". Then, working backwards, if .010" seemed too much, the designers might feel that a lower figure would be necessary, say (generously, they hope) of .0075" either way. Then they would divide by 5 and assign .0015" to each individual or three-fourths of the original figure of .002". In other words, the individual tolerances are now tightened in order to achieve an overall goal.

By the laws of probability, it would be an extreme rarity to bring together all 5 items of the assembly at the high or at the low of their respective distributions. In random assembly, we would expect the five pieces to represent all degrees of their respective distributions, some high, some low, but mostly near their mid-points.

For example, if one in every 1000 units were on the high tolerance, the chance of getting five of them together would be $(1/1000)^5$ or 1 in 1,000,000,000,000,000 (one thousand million million) assemblies - not a very probable occurrence.

The real problem is not the chance of these five getting together, but of finding the probable practical spread of random assemblies of 5 items, or what spread would contain 99 plus percent of all the assemblies. Roughly this would proceed as follows:

The only data we have is the $\pm$ .002" or .004" spread allowed units of each kind. If the normal distribution just "filled" the tolerance (the worst possible condition) this would be equivalent to 6 sigma which Dr. Edwin Olds, who has done much to study this condition (3), calls the "natural tolerance". He then sets up the formula:

$$t_s = \sqrt{t_1^2 + t_2^2 + t_3^2 + t_4^2 + t_5^2}$$

Since $t_1 = t_2 = t_3 = t_4 = t_5 = .004$

$$t_s = \sqrt{5(.004)^2} = \sqrt{.000080} = .00894$$

Thus, in practice, the maximum spread of the 5-item assembly would be .00894 or .00457" on a side, well within the .0075" either way which was set as a working assembly tolerance.

Now that we have the camel's nose inside the tent, let's go in a little farther. The specification writer conceded that .0075" would be enough. We might see how wrong it was to set .0015" for the individuals. We simply set up the formula to let X equal the individual 6 sigma limit, and solve for it.

$$t_s = \sqrt{5X^2}$$

$$t_s^2 = 5X^2$$

$$X^2 = t_s^2/5 \text{ with } t_s \text{ now set at } .0075$$

$$X^2 = .000045$$

$$X = .006708$$

$$X/_2 = .003354$$

We could have set the individual item tolerances at .003354" either way (or round it off to .003") and be quite certain that no trouble would occur in assembly. In other words, the tolerance could be made half again larger.

The Producer

Now let us turn to the producer and a very few of the many applications he can use. For one thing, the producer has a big advantage. Where the consumer could only look at the product as a whole (as a grouped population in a shipment, for example) the producer can break up his study into sub-groups taken during production. This has one big advantage - he can remove for detailed examination only that portion of a production which needs it.

Five problems will be illustrated:

1. Correlation of Specifications.

Sometimes a change in the method of expressing a specification is desirable during production (1). An example is given for the

capacity of a glass container. It may be desirable to the consumer to have the capacity for the content he puts into the bottle to center on a point in the neck of the bottle. The producer, because he controls capacity by the difference between mold displacement and volume of glass per bottle, may more easily use the complete or overflow capacity of the bottle. He must then have some correlation between the two specifications so that he can control one by measuring the other. The method used is known as a correlation analysis.

Figure 18 is a graphical portrayal of the method. Each dot on the chart is for a single bottle. Each dot at the same time shows the overflow capacity and the height on the side of the bottle to which the stated contents would come. The several hundred dots on the chart represent the measurement of that many bottles from several manufacturers. The correlation is good.

One usually starts with a table, giving for each bottle the two measures made, called "paired observations". It usually pays also to plot the data at this point because if the graph shows an obvious lack of correlation, the calculation step need not be taken. Usually however the calculations are carried out and result in the formula for a straight line which gives the average correlation together with the outside limits of the variation.

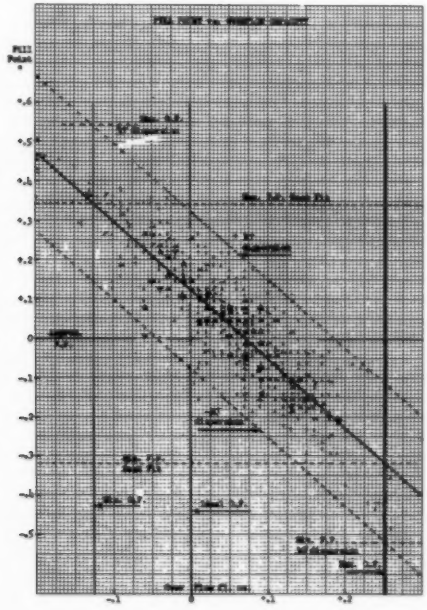


Fig. 18

In this case a simple transition answered the specification question. The fillpoint specification had been "guessed" too low by about an eighth of an inch. The "zero" or ideal of the overflow was equivalent to that amount from the assumed filling point ideal on the scale to the left. The filling point equivalent to -1/8 oz. overflow (the minimum tolerance) was +.34", and that equivalent to the maximum overflow of +1/4 oz. was at -.32", both measured at the assumed ideal fillpoint.

This process then gave the producer a working specification in one term which could be translated during production to the language of another specification.

2. Tool Wear

One common mistake in specification writing is to measure a few items and then set a specification which assumes no further variation. Where this is most dangerous is in operations such as machine tooling where the tool gradually wears during production. Every time a tool must be removed for grinding, there is loss of production until the new tool is properly set. This is an economic loss which must balance with the specification limit. If limits are too close, the constant re-setting of tools would make the cost prohibitive.

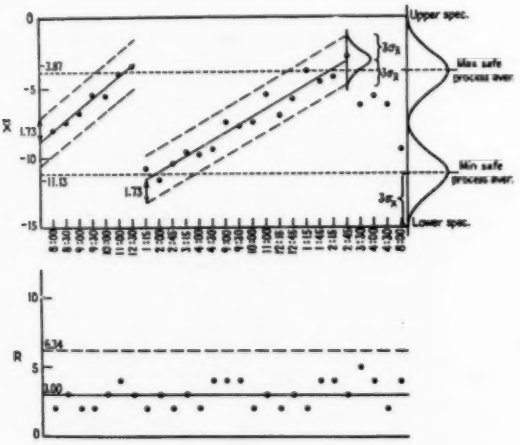


Fig. 19

By permission from "Engineering Statistics and Quality Control"
By: Irving W. Burr
Copyright, 1953. McGraw-Hill Book Company, Inc.

What is needed therefore, is a specification which would allow moderate wear. At the same time some system of warning must be given as to approach to the specification limit.

The result is a control chart with slanting lines for averages and limits, such as I show here from Prof. Irving Burr's "Engineering Statistics and Quality Control", P. 269 (4).

The spread of the averages of samples of 5 consecutive units is shown as the vertical distance between the dotted lines. The true (probable) spread of individuals would be shown by one of the heavy normal curves to the right, filling only about half the specification distance. Consider these two heavy line curves as the beginning and ending positions with many positions in between.

This graph shows a fairly workable condition of specification. The specification spread is twice the spread from any one time. Tools would last about two hours. If the specification were cut in half, tools could be used about 10 or 15 minutes - an impossible condition economically.

3. Effect of reducing a specification tolerance

This leads to a question as to what would happen if a specification tolerance were reduced. Our Chart 20 gives a graphical analysis.

Suppose the total tolerance, on some scale, were 140 units. Suppose the spread at any one time of production was 40 of these units. (Ignore for the moment the cry "you don't need all that tolerance".)

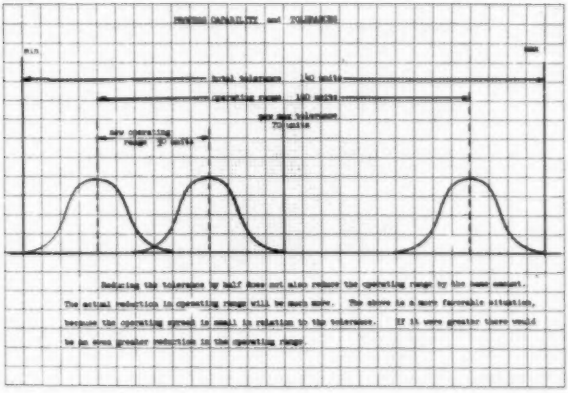


Fig. 20

Perfectly controlled production would mean that it started with the 3 sigma tail of the distribution curve just touching the left hand tolerance and continued until it crossed the graph and the opposite tail just touched the opposite tolerance. This could just as well be shown by the centers of the two curves. The operating range would be 100 units.

Now suppose that the tolerance were cut in half. The secondary position of the curve is now moved so that the 3 sigma tail touches the new tolerance only 70 units away from the first. The center distance is only 30 units. Halving the tolerance has cut the operating range by 70%.

The illustration only gives a tolerable condition for tool and mold wear situation. Each specific problem depends on the relation between the sigmas and the tolerance.

4. Control Charts

Much mention has been made of control charts, and Chart 21 gives a typical chart. Its function is to help keep production on a steady keel by analyzing for trends and for causes of deviation not inherent in the process itself.

The control chart, as such, does not show a specification limit. Specifications are for individuals - control charts deal in the averages of small samples. Some means must be provided to tie the two together.

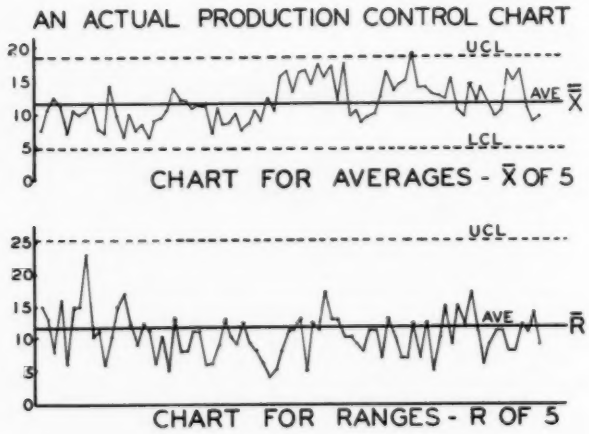


Fig. 21

The several methods of doing this result in what is known as "modified" control limits. Either the specification is changed (for working purposes only) into a specification for the average of the sample size, or a secondary line is placed on the control chart giving a limiting position of the average of the sample, just as for tool wear. In either event this is accomplished by subtracting an amount from the specification maximum, or adding to the minimum. Usually this is approximately the difference between 3 sigma of the population and 3 sigma of the average of the sample.

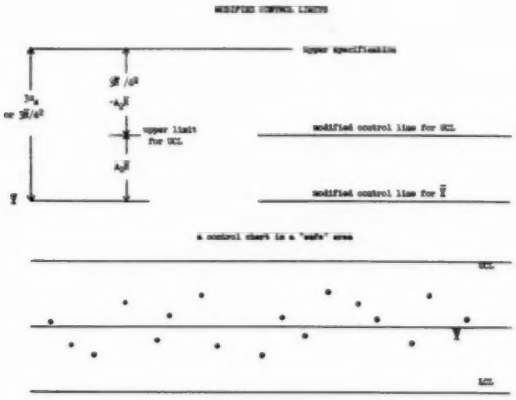


Fig. 22

5. Process Capability

My last point is most important of all - the study of process capability in relation to specification. It is so important that any kind of complete treatment is too much for this paper. Fundamentally, it is a study of the several phases of variation in production with the view of finding if it will meet specifications. It includes a study of the following:

Variation of the individual sampling.
- Instrument error
- Sampling error
- Inspector variation

Variation of the product at any one moment of time.

Variation of the product from one time of production to another time.

Variation between multiple or parallel machines.

Variation from one production run to another.

One method of handling this has become known as the SPAN capability method, which takes its initials from the words "Systematic Procedure for Attaining Necessary capability". It is a statistical procedure, not too complex, but definitely systematic which can be applied. (5)

Many of you are possibly "lost", statistically, by this time, even though I have purposely avoided details, each of which would have been a discussion in itself. You may be inclined to "skip the whole idea". Such inertia is one of the problems of today - it leads to lack of understanding and to futile disagreement. Let me urge you to investigate the statistical method of "finding the facts before setting the specifications". Its use will save thousands of dollars of unnecessary loss - your money and others' money.

Summing up this paper, the correctness of a specification is determined by the mathematical relationships between use and production, together with cost. Specifications fundamentally should be based on fact. The facts must be known first. They must come from a study of the product, of its use, and of its production. One of these three considerations alone, or any pair, is not enough. Therefore, correct specifications mean an open, cooperative effort by maker and consumer. Each must be conversant with the problems of the other and both must be willing to study the overall problem.

Finally, "Specifications are not rubber bands - to be stretched when necessary, then a little wider, and a little wider, until the whole structure falls".

REFERENCES

1. Toulouse, J. H., "Finding the Facts Before Setting the Specification", Fifth Annual Convention, A.S.Q.C., 1951

2. Toulouse, J.H., "Load Concentration in Corrugated Paper", A.S.T.M. Bulletin No. 152, May, 1948

3. Olds, E.G., "Practical Tolerances", Michigan Forum, A.S.Q.C., 1952

4. Burr, Irving W., "Engineering Statistics and Quality Control", McGraw-Hill Publishing Company, 1953

5. Seder, Leonard A., and Cowan, David, "The SPAN PLAN - A New Method for Process Capability Analysis", 7th Annual Convention, A.S.Q.C.

STATISTICAL QUALITY CONTROL IN AUTOMOTIVE STAMPING PRODUCTION

J. Mehalek
The Budd Company

When I took on the task of determining what was being done in the application of "Statistical Quality Control in Automotive Stamping Production" I knew how we controlled quality at Budd's. I also knew how it was done in some of the other automotive manufacturing plants, but I did not know what the entire industry was doing.

In the more than fifty years and three major conflicts that the Automotive Industry have experienced, they proceeded to grow without any real major setbacks, so that today they stand collectively as the largest single manufacturing industry of a consumer product. It only stands to reason that under these conditions they must have good sound systems of quality control.

In studying these systems of quality control still referred to in the major portion of the industry as "inspection", it was learned that in a number of cases a method had developed that is peculiar to that industry in the stamping of automotive parts. This system, based on the thought of the older school, that if a job is:

(a) Well designed
(b) Properly engineered
(c) Correctly planned
(d) Satisfactorily constructed
(e) And initially checked and O.K.'d
(f) Inspection isn't necessary.

It can be readily seen that this essentially is a form of control - control of the functions that can possibly cause difficulties, thereby preventing such difficulties. Although this did not always work, it is surprising to see how often it did, and can be credited with the earlier successes of the industry where quality and production were concerned. As with all industries, as the product became more complex and production problems more demanding, the system of quality control generated into a system of inspection and re-work, and finally quality control with the use of records and statistics, although it is not referred to in this manner.

Fifteen of the major automobile manufacturers and automotive stamping plants were contacted for information in regard to their application of Statistical Quality Control in automotive stampings. Returns were received from seven. To avoid comparisons of either the response to the questionaire or methods of application, no reference will be made to these companies to distinguish one from the other, than to list their names for assistance credit at the end of the paper.

The information requested was as follows:

1. Initial application - date
2. Type of application - product
3. Whether application directly controlled press operations,

and/or product of same.
4. Method of charting and control
5. Results and summary
6. Copies of charts
7. Pictures of operations, if possible

You can readily see if there had been 100% response what a beautiful statistical report could have been given on the subject. Without it we will use the information received to give you some idea of what is being done and how.

By turning the clock back a few years we can see just how the present methods of quality control evolved. In going back, many names were encountered of those who had pioneered in the early applications of inspection and quality control in the automotive stamping field.

Previous to the mid-thirties, detailed drawings on auto stampings were accepted by the manufacturers and immediately used in making wood models which in turn were used in making dies, templates and other tools. These models were seldom checked by the Inspection Department. The result was that any error in the model became an error in the die or fixture built to that model, resulting in stampings that were wrong and would not assemble properly. About 1933 this was changed so that all models were checked by the Inspection Department before releasing for die or template building, or any other use.

About the same time lists of complaints were published and distributed to the shops, with hourly checks being made on the material coming off the lines to see if corrections were made. This was not exactly accepted by the production people with open arms, but when understood and results shown, they went along with the idea. This led to the next step - the "Check Off System" which has proved quite satisfactory through the years, being continued in use today in an improved form. The "Check Off System" uses a form covering a major part, upon which is detailed all the important dimensions with provisions for indicating the condition of any particular dimension. These reports were then used in the acceptance or rejection of the material in question. They were then examined by the person in charge of inspection who would initiate the necessary action to correct the existing conditions.

Although even today charts do not predominate in the automotive stamping operations, facts are so organized as to give the various key personnel (Managers, Superintendents, Chief Inspectors) a continuous comprehensive picture of conditions and quality standards as they exist, in a manner to meet any requirements of the management. Statistics are compiled on acceptances, charges, costs, scrap, etc. being fed into automatic recording calculators and broken down in such manner as to be available for managerial purposes.

To better appreciate the Quality Control Function of an Automotive Stamping Plant, one must understand the problems peculiar to this method of manufacture and the means used for its control.

There are two distinct variables to be contended with in the

manufacture of stampings:

1. The sudden change of quality which wou
in a histogram by a sudden shift of th
to one side or the other. This could
ken tool, such as punch, shear edge of

2. Then there is the gradual change of qu
sult in the broadening of the base of
of the histogram. This being due to n
tools and would be so very gradual it
able on any normal run and perhaps not
of the job. The exception being on heavy gauge automotive structurals such as chassis frames, where radii washout more rapidly due to the tremendous forces required for forming these parts. One condition changes so rapidly and the other so slowly that the usual statistical methods of an "$\bar{X}$" and "R" chart or even the "P chart can be of little use after an initial survey.

This does not mean that there is no reliability program in effect in press shops. The various companies have their own methods of approach to this problem, usually done on a sampling basis. One of the cooperating companies use a "Check Off Sheet" for major stampings and sub-assemblies and developed a beautiful and effective "Bar Chart" indicating a very decided improvement in their product.

We can use the following as an example of a reliability program that can be considered as in general use with automotive stamping plants.

After die try-out and a detailed first piece inspection of the stamping resulting in acceptance and approval to start the run, periodic checks of the stamping are made at regular intervals. This may be done with jigs and fixtures or just templates, depending on the requirements. The information gathered by this check is recorded on charts or check sheets (similar to those mentioned before, detailing all the major dimensions) indicating the progress of the job. From past experience changes are so gradual that a run is usually completed before any appreciable change takes place. In fact, a number of runs or the entire job may be completed without any appreciable change.

Where there is an indication of trouble such as the constant breaking or wearing of punches or the washing of a radius, that can be pinned down to a definite cycle, a study is made by charting the time cycle to determine the best time for corrective action to minimize time loss due to maintenance. When possible this change will be made during the nearest break period to the end of the cycle, so that there is no interference to production or reduction in the standards of quality. Prior to this system, changes were only made after the creation of defective material, which entailed sorting out of the defective parts and subsequent re-work. This simple control method eliminated the additional expense of time and manpower for corrective action.

There are times when an inspection check of the stamping may

and/or product of same.
4. Method of charting and control
5. Results and summary
6. Copies of charts
7. Pictures of operations, if possible

You can readily see if there had been 100% response what a beautiful statistical report could have been given on the subject. Without it we will use the information received to give you some idea of what is being done and how.

By turning the clock back a few years we can see just how the present methods of quality control evolved. In going back, many names were encountered of those who had pioneered in the early applications of inspection and quality control in the automotive stamping field.

Previous to the mid-thirties, detailed drawings on auto stampings were accepted by the manufacturers and immediately used in making wood models which in turn were used in making dies, templates and other tools. These models were seldom checked by the Inspection Department. The result was that any error in the model became an error in the die or fixture built to that model, resulting in stampings that were wrong and would not assemble properly. About 1933 this was changed so that all models were checked by the Inspection Department before releasing for die or template building, or any other use.

About the same time lists of complaints were published and distributed to the shops, with hourly checks being made on the material coming off the lines to see if corrections were made. This was not exactly accepted by the production people with open arms, but when understood and results shown, they went along with the idea. This led to the next step - the "Check Off System" which has proved quite satisfactory through the years, being continued in use today in an improved form. The "Check Off System" uses a form covering a major part, upon which is detailed all the important dimensions with provisions for indicating the condition of any particular dimension. These reports were then used in the acceptance or rejection of the material in question. They were then examined by the person in charge of inspection who would initiate the necessary action to correct the existing conditions.

Although even today charts do not predominate in the automotive stamping operations, facts are so organized as to give the various key personnel (Managers, Superintendents, Chief Inspectors) a continuous comprehensive picture of conditions and quality standards as they exist, in a manner to meet any requirements of the management. Statistics are compiled on acceptances, charges, costs, scrap, etc. being fed into automatic recording calculators and broken down in such manner as to be available for managerial purposes.

To better appreciate the Quality Control Function of an Automotive Stamping Plant, one must understand the problems peculiar to this method of manufacture and the means used for its control.

There are two distinct variables to be contended with in the

manufacture of stampings:

1. The sudden change of quality which would be illustrated in a histogram by a sudden shift of the bell shaped curve to one side or the other. This could be caused by a broken tool, such as punch, shear edge of a blanking die, etc.

2. Then there is the gradual change of quality that would result in the broadening of the base of the bell shaped curve of the histogram. This being due to normal wear of the tools and would be so very gradual it would not be noticeable on any normal run and perhaps not even during the life of the job. The exception being on heavy gauge automotive structurals such as chassis frames, where radii washout more rapidly due to the tremendous forces required for forming these parts. One condition changes so rapidly and the other so slowly that the usual statistical methods of an "X" and "R" chart or even the "P chart can be of little use after an initial survey.

This does not mean that there is no reliability program in effect in press shops. The various companies have their own methods of approach to this problem, usually done on a sampling basis. One of the cooperating companies use a "Check Off Sheet" for major stampings and sub-assemblies and developed a beautiful and effective "Bar Chart" indicating a very decided improvement in their product.

We can use the following as an example of a reliability program that can be considered as in general use with automotive stamping plants.

After die try-out and a detailed first piece inspection of the stamping resulting in acceptance and approval to start the run, periodic checks of the stamping are made at regular intervals. This may be done with jigs and fixtures or just templates, depending on the requirements. The information gathered by this check is recorded on charts or check sheets (similar to those mentioned before, detailing all the major dimensions) indicating the progress of the job. From past experience changes are so gradual that a run is usually completed before any appreciable change takes place. In fact, a number of runs or the entire job may be completed without any appreciable change.

Where there is an indication of trouble such as the constant breaking or wearing of punches or the washing of a radius, that can be pinned down to a definite cycle, a study is made by charting the time cycle to determine the best time for corrective action to minimize time loss due to maintenance. When possible this change will be made during the nearest break period to the end of the cycle, so that there is no interference to production or reduction in the standards of quality. Prior to this system, changes were only made after the creation of defective material, which entailed sorting out of the defective parts and subsequent re-work. This simple control method eliminated the additional expense of time and manpower for corrective action.

There are times when an inspection check of the stamping may

not reveal some of the details of difficulties to be encountered in assembling. This may be due to stampings being made to opposite extremes of tolerance, metal springback or just the close tolerances to which it is necessary to make the part. As the Automobile Industry is built around mass production, which in turn requires ease of assembly, considerable concern is shown when this condition takes place.

To cope with this situation the automobile stamping industry have designed special inspection and checking fixtures. These at times become very complicated and may accomodate the parts for an entire body or chassis in such manner that the parts can be checked in their relative car positions. This also simulates assembly operations and proves the assembly of the parts. Again your "Check Sheet" comes into play to record the result of an inspection check. These inspection checks are made at regular intervals, the results being recorded and plotted to indicate any trends. When there are indications toward deviations, that if permitted to continue would be cause for rejection or cause difficulty in assembly, corrective action is taken. These corrections are usually made during a down period to keep production losses to a minimum, which makes it so vital to be able to anticipate any possible deviations. When a sudden deviation takes place, due to a tool breakdown, immediate action is necessary to correct the condition. If this is not possible, these parts are put aside and corrected. You can rest assured that this isn't permitted to happen very often, as it would greatly increase the cost of the product. These controls are very effective in keeping such incidents to a minimum.

When one stops to consider the number of mating parts and matching holes that in turn must be matched or mated with parts possibly from other sources of supply, such as other shops or companies, your quality control must be good. This ability to accomplish this control was not attained by simple inspection procedures, but by the use of endless records, otherwise called statistics. Accumulating and organizing these facts to realize the potentialities of a manufacturing process, to know how close these parts may be held. Naturally, this being done not alone by quality control, but with the aid of Engineering Planning, Tooling and last but not least, the Press Shop Personnel. To realize the effectiveness of this program one only has to look back a few years when the permissable tolerance in stampings was 1/32" and sometimes more. Today it isn't rare to have these tolerances held to 1/64" and even less.

The quality control of automotive stampings does not end in the press shop but carries on to the assembly floor where a very close surveillance is kept over the condition of stampings in relation to their ease of assembly. Reports on assembly conditions giving difficulty are directed back to the press shop, where the tools are carefully re-checked and the parts are again checked to the piece drawing. When the tools and the parts check satisfactorily and there is no explanation for the trouble in assembly, Engineering is called into the picture to make a study of the problem, effecting an engineering change, should it be necessary, to correct the difficulty.

The problems of the quality control function in assembly differ

a great deal from those of the press shop. It is in assembly where the most work has been done in the statistical quality control field. With the large number of manual operations in assembling, such as welding and finishing, the necessity for non-destructive testing, and the checking of such intangibles as metal finish, have provided a fertile field for sampling techniques, control charts, quality rating factors and other devices permitting the use of pyschological and visual aides. They have in most cases proved satisfactory as indicated by the papers and publicity covering the subject, but this is another story and should be handled as a separate presentation.

Summarizing this paper on the application of Statistical Quality Control in Automotive Stamping Production, let us see what we have.

Generally speaking, statistics are used in the application of quality control in press shops, but not in the sense, as statistical quality control people recognize these applications.

We may ask ourselves why?

Can it be that the potential of application is limited due to:

1. Material control handled by other than the Press Shop, such as, the laboratory or receiving inspection.

2. Tools, nature of them reduces variables to a minimum, requiring little attention after initial check and acceptance; wear being negligible in the normal application of those tools.

3. Inspection Controls

   (a) Detailed initial inspection providing for initial close control, assuring satisfactory tooling.

   (b) Periodic sampling, with use of intricate gauges and fixtures, simulating end use of the product, plus report of results.

   (c) Liaison system between press shop and assembly floor, assuring satisfactory application of the parts.

Can we say that Management has not been sold in the use of Statistical Quality Control in press shops, when we know its use has been adopted in assembly operations and has been used on occasion, in the production of stampings.

These questions and others can probably be best answered by you, particularly those of you who have a direct association with automotive press shop operations.

QUALITY CONTROL IN THE MANUFACTURE OF FILM COATED MAGNET WIRE

W. E. Bramer
John A. Roebling's Sons Corporation

Before we begin to discuss the use of Quality Control Techniques in the manufacture of Film Coated Magnet Wire let us take a few minutes to look at the product in general. The Electronic Industry and the motor and transformer industries have progressed so fast, during the past decade, that the quality requirements for this Magnet Wire are now far in excess of those necessary a few years ago. For example - the manufacturer in these fields is required to wind many turns of magnet wire into a highly restricted area; therefore the overall diameter and the inherent size variability of the wire become increasingly important. Again, the manufacturer is continually required to produce smaller and smaller equipment and still keep his quality ratings at the same level. His ability to do this depends largely on his ability to obtain high dielectric strength magnet wire; in fact, for this particular end use, a wire with a dielectric strength far in excess of specification requirements of a few years ago is a prerequisite.

In order to produce a finished product with a quality level well in excess of general specification requirements, it is necessary to start with the raw material and to control the processes carefully at each manufacturing point, as the transition from incoming raw material to outgoing finished material progresses.

One of our first applications of Statistical Quality Control Techniques at John A. Roebling's Sons Corporation was to control the manufacture of Hot Rolled Copper Rods. This control was established about ten years ago and has since been extended to cover the entire manufacture of Magnet Wire. We are happy to be able to share our experiences in this field with you.

The raw material from which Magnet Wire is made first enters our plant in the form of "Copper Wire Bar", which is purchased from several reputable copper refineries, and is subjected to a routine visual inspection before being released for production. The copper wire bar is loaded on special conveyors and delivered to the "Rod Mill" where it is hot rolled into "Copper Rods" of the required diameter (usually 5/16" or 3/8") in coils of 24" I.D., weighing about 250 lbs. each. Upon delivery to the "Wire Mill" these coils are first cleaned in a sulphuric acid solution to remove all scale and then dipped in a soap solution to aid in the wire drawing operation. The cleaned rod is first shaved, to remove any surface imperfections, then cold drawn to the required size (annealing in process if required) and sent to the Enamel Shop where it is covered with successive coats of enamel or varnish as the order may specify and placed on reels, spools, or in drums for shipment.

So much for a general review of the manufacturing process. Now let us review it more slowly and see how Quality Control Techniques help, at strategic processing spots, to produce a finished material that surpasses specification requirements.

INCOMING INSPECTION - The incoming inspection of the Copper Wire Bar is quite routine. The bars are examined visually for surface

imperfections and pouring defects, particularly high rolled edges that would be injurious to the rod surface after hot rolling. Records are kept and the Quality Level of the hot rolled rods produced from wire bar from the various refineries are compared with the incoming inspection findings.

ROD MILL - The transition from copper wire bar to hot rolled copper rods is accomplished by heating the copper wire bar to approximately 1700°F in a long furnace, under a reducing atmosphere, and then hot rolling it through a number of rolling passes. This operation reduces it in cross section and elongates it in length by a series of oval, diamond, and square shaped passes to the final round rolling pass, called the finishing pass, which produces the hot rolled rod of the size required.

For a number of years prior to the introduction of Quality Control Techniques into our Rod Mill we used an inspection procedure whereby we took samples from each of the four finishing reels at specified intervals, and after twisting and examining the samples, graded them into class one, two, or three depending on the number and degree of imperfections noted. An attempt was made to use only class one rods for Magnet Wire end use, class two for high grade strands and fine wire, and to divert class three rods into line wire or other non-critical end uses.

In 1948 the amount of class three rod produced began to increase alarmingly and by June of that year was averaging nearly 25%. About that time we were becoming seriously interested in starting a Quality Control program but did not know just where to begin. The Hot Rolled Rod situation was becoming quite serious, so: that became the starting point.

Using dead data, and for the moment calling class three rods defective material, we worked out a Percent Defective Chart on a daily basis and established control limits around the monthly average of the daily percent defective. We fully realized that we were taking certain liberties with pure statistics, in as much as the daily lot sizes were different and we were ignoring this difference. Our main purpose was the correction of a bad situation and from this simple beginning we have learned that the simplest approach, not always statistically pure but at least understood by the rank and file, often pays the highest quality dividend. Our next step was to discuss this chart thoroughly with the operating personnel in the Rod Mill and to erect a large control chart, which was posted daily, where everyone connected with the Rod Mill could see it. Daily percent defectives which were within the control limits established from the operation of the previous month were posted in green. Percent defectives falling above the upper control limit were posted in red and those below the lower control limit were posted in yellow. As regards to the use of the chart, the operating personnel were requested to do only three things. First, if the daily percent defective fell within the established control limits, regardless of the level, they were to continue operations with no changes. Second, if the daily percent defective fell above the upper control limit, they were to do all in their power to locate the cause of the trouble and to correct it. Third, if the daily percent defective fell below the lower control limit, they were to attempt to determine the reason for the exceptionally good run and to make such

processing changes as might be necessary to take advantage of it in the future. They were assured that if this plan was followed faithfully the percent defective would be lowered each month until it reached a state of control which would represent the inherent quality level of that particular mill and operators. By October the percent defective had dropped to 8% and later to about 2-1/2% where it has remained. As a result of this, the quality level of the mill was raised to the point that it was no longer necessary to segregate the rods for various end uses. The entire personnel in the Rod Mill became very quality conscious and other manufacturing units, which had been following the procedure with interest, began to request Quality Control programs of their own.

WIRE MILL - There are, of course, many problems which present themselves in a wire mill. If we concentrate on Magnet Wire we find that the surface and gauge of the finished wire are both of prime importance. Assuming good wire mill practice, the surface is largely dependent on the quality of the hot rolled rods from which the wire is drawn and we have already seen how Quality Control can be used to advantage to hold the quality level of this product at a high level. The diameter or gauge of the finished wire is also vitally important and we would like to share with you our experiences in the use of very simple Quality Control techniques to control the diameter of this wire within limits more restrictive than specification requirements.

We were well aware that certain customers for Magnet Wire were desirous of obtaining a wire with a closer gauge tolerance than general product specification limits. We were also alert to the economic fact that when a product which is sold by the foot, is made from raw material which is purchased by the pound, it is to the manufacturer's financial advantage to produce as many feet of finished product per pound of raw material as possible. While this does not pertain to Magnet Wire it is true of many products in the insulated conductor field, and all producers realize the advantage to be had in holding the gauge of individual wires and strands as close to specification minimum as possible, without suffering costly rejections by violating these minimum requirements.

The question was asked many times in our organization, "Can we produce a copper wire with only one half the variation in diameter as that allowed by industry specifications?" Some answered this question in the affirmative and some in the negative, but all were guessing to a large extent. Their answers were based on their individual opinions of what could be done, but were not founded on the facts of the case.

Our second step in this particular Quality Control program was to determine just how close the diameter of the basic copper wire could be held, using the present equipment, without unduly increasing the manufacturing cost.

A process capability study was made on a small group of four wire-drawing machines, drawing finished wire, to determine the degree of variability inherent in the process as regards to the gauge of the wire produced. As in the case of the Rod Mill this was done using the simplest of Quality Control techniques - in this case the average and range chart. Our approach was to take a sample of wire from each of the four machines at regular intervals, gauge the samples, average the

diameters, and plot the average and range on a control chart. In a relatively short time it was learned that the equipment was entirely capable of producing wire within industry specification limits, but was incapable of holding it to one half of these tolerances. We extended our experimental work and studied samples taken from individual machines in the group, instead of from the group as a whole, and found that while some machines were capable of maintaining the desired tolerance, others were not. In addition, we were quite surprised to learn that one of the machines was producing wire which was not entirely within current specification limits and was responsible for some of the finished wire rejections we had been experiencing.

From this short and inexpensive study it was learned that: 1. The present process, assuming a slight modification on the one machine, was capable of maintaining industry tolerances; 2. The present process was not capable of producing wire to the doubly restrictive tolerance desired; and 3. There was a good indication that if further investigation was made and additional expense incurred in refining the process it would be quite probable that these restrictive tolerances could be maintained.

Following this process capability study a meeting was held at which it was decided to modify the process slightly and to hold the total product within the lower three quarters of the existing specification. Average and range charts were immediately set up on this product and we began producing wire with an average diameter in the lower half of the existing specifications and the upper and lower controls set so that no wire would fall below the minimum specification and only a very small percentage would fall above the commercial specification nominal.

The decision made at that meeting was based on fact, not someones opinion, and was proven by the uniformity of the product produced from that time on. As a result of this, we have not only lowered manufacturing costs, by reducing rejects and obtaining more conductor footage per pound of copper on many items, but at the same time have supplied our Magnet Wire customers with a product which is more completely designed to their special requirements.

Thus far in our discussion, all that has been discussed could well be said about copper wire in general, drawn for any end use. The material destined for the Enamel Shop is finished on reels or spools of appropriate size. These spools or reels are paper wrapped and packed on skids for delivery to the next and final processing operation.

ENAMEL SHOP - Let us take a few moments to look at a typical Enamel Shop operation. Basically the main purpose of the process is to apply a thin, smooth film of insulating varnish or enamel, as the case may be, to the surface of the copper wire. Normally the dielectric strength is proportionate to the thickness of film and both are of vital importance to the user. The insulating film is applied by repeated passes of the wire thru a varnish applicator. Dies control the thickness of the coating. The coated conductor passes thru the baking oven where the enamel or varnish is thoroughly cured and returns to the applicator for the required number of successive coats.

There are three basic sizes of equipment which we have in general use today to cover the wide field of sizes used. We refer to the large

enameling machine and baking oven, generally known as the "M Machine", which normally handles sizes ranging from #8 to #23 A.W.G. We next have the intermediate group, known as the "I Machines", handling sizes from #23 to #32 A.W.G. and also the fine wire group, better known as the "F Machines", which are capable of handling sizes from #32 to #44 A.W.G. The basic operation of the three machine groups is the same, consisting of alternating cycles of coating and baking. The "M" and the "I" machines operate vertically, and the thickness of film is controlled by sizing dies. The "F" machines operate horizontally and the thickness of film is controlled by a rotating rod which has been grooved to permit the proper amount of varnish to flow on the wire, and by a close control of the viscosity of the enamel or varnish. The "M" and "I" type machines are equipped with a high temperature, neutral atmosphere, annealing chamber which softens the wire prior to the coating cycle, whereas; the "F" type machine depends solely on the heat in the baking oven for the softening effect.

In connection with the final stage in the manufacture of Film Coated Magnet Wire there are several very important characteristics which must be watched carefully and controlled within close limits. Statistical Quality Control has proven to be of great value in maintaining the quality level of this product on a high standard.

The vitally important characteristics of Magnet Wire fall into two categories: vis, Physical and Electrical.

PHYSICAL CHARACTERISTICS

1. Gauge

   a. Conductor size – generally referred to as the bare wire diameter, is vitally important. Too small a conductor size may be responsible for high rejections on the part of the coil winder, due to high electrical resistance; and conversely, if the diameter of the conductor is on the high side of the size tolerance the coil winder may not be able to wind the required number of turns of wire into the highly restricted space allowed him by the coil designer or user.

   b. Thickness of film – normally referred to as insulation buildup, is also vitally important. A wire with minimum insulation buildup may well meet general industry specification, but a coil wound from this wire may fail under an electrical voltage surge test applied to the finished coil or motor. On the other hand too high an insulation buildup may render it impossible for the coil winder to hold his finished product within required size restrictions.

   c. Overall Diameter – commonly referred to as O.D., is equally important as it is a combination of the above mentioned dimensions and is of vital importance to the coil winder in maintaining finished coil dimensions within size restrictions.

2. Softness – We all recognize copper to be one of the softer metals. Strangely enough the degree of softness plays a very

important part in the use of magnet wire in coil winding. The coils produced by the coil winder are not all round and cylindrical. Some are square and some rectangle. Some are wound as skeins, similar to yarn, and then positioned into coils. Some are quite sharp at the corners and some are nicely rounded. All this variability in coil size and shape imposes a definite degree-of-softness requirement on the wire from which the coil is made. The magnet wire must be soft enough to conform closely to the core, or basic shape of the coil, under many varying winding conditions of speed and tension. It must be capable of withstanding bending at high speed around sharp angles and stay put with little or no spring back, for as it can be readily seen, the accumulated spring back from numerous layers of coil winding would result in mushy, oversized coils unsuitable for their required end use. On the other hand, if the magnet wire is too soft, any unusual increase in the winding tension during coil winding will place the wire under a strain greater than the elastic limit of the wire and the diameter may be reduced, pulled down as we call it in the mill, to a point where the electrical resistance of the finished coil may be raised beyond specification limits. Modern high speed winding equipment has imposed more critical demands on the magnet wire producer in the past few years than ever before. These machines are extremely intricate and pass the wire over numerous small sheaves and guides to the coiling head thus requiring a uniformly soft and well spooled wire for efficient operation. Most users of magnet wire stress their requirement for extremely soft wire, however the manufacturer who knows his product, and is conversant with the use to which it will be put, will produce a wire that is neither too hard nor too soft, but just right for the individual customers end use.

ELECTRICAL CHARACTERISTICS

1. Dielectric Strength - The ability of the insulating film on the magnet wire to withstand voltages far above specification requirements, in the form of surge voltages, is highly important to the coil winder. All other things being equal the dielectric strength is in proportion to the thickness of insulating film. We have already learned of the importance of the space factor which makes it advisable to hold diameter and film thickness to a minimum; therefore, the processing must be adequately controlled so that the maximum dielectric strength is maintained with a minimum film thickness.

2. Continuity - The complete envelopment of the wire by the insulating film, with no pinholes, voids, cracks or openings which might cause an electrical short in the finished product, is of vital importance to the coil winder. Industry specifications allow a maximum of 15 pinholes or discontinuities per one hundred feet of length. If the manufacture is to enjoy a carefree existence, if such is possible, the process must so be controlled that the coil winder will receive material with a minimum of discontinuities throughout the entire length.

The requirements of the industry specifications referred to thus far lend themselves nicely to Statistical Quality Control Techniques and

we have profited greatly from their use. There are also mechanical, thermal and chemical requirements of the specification, such as tests for solubility, aging, adhesion, etc., which do not lend themselves so readily to these techniques. These characteristics are dependent on the raw materials used and on the basic manufacturing process. Once the raw material and the process standard are established and are compatible, there is little variability encountered and the process, in this respect, can be controlled by small sample sizes taken at extended intervals.

In setting up a Quality Control program in the Enamel Shop the first problem was to work out a suitable sampling plan. As it is frequently necessary to run more than one size on a machine at a time, the unit lot, for sampling purposes, was designated as a particular size running on a given machine during a specified period of time. The plan decided upon was to take four samples from every machine in operation and if there was more than one size or grade of magnet wire running on a particular machine, four samples were taken from each different size or grade. The samples were taken at specified time intervals, by the process inspectors working on the floor, and were sent to the tester working in the control laboratory. Upon receiving the samples the testers responsibility is to test all samples received for:

1. Gauge
   a. Overall diameter
   b. Bare copper diameter
   c. Insulation buildup (by subtraction)

2. Softness

3. Dielectric Strength

4. Continuity (if size requires)

The results of these tests are then analysed and plotted statistically, on simple average and range charts, and the results are made available immediately to the Production Department to enable them to control the process satisfactorily. An individual control chart is kept for each size and grade of wire manufactured and, although a particular size has not been made for some time, the previous chart will be removed from the file, and added to, when that size next appears on the production schedule. In this manner we have a continuous quality record of all wire manufactured, by size and grade, showing the quality level of each manufacturing run with relation to any other run of like material and also a record of the quality level from day to day within any production run. As the date of manufacture appears on all reels, spools and shipping containers it is possible to check the quality level at which the particular item was produced, even after delivery to a warehouse or customers plant. In addition to the testing referred to above for control chart purposes, the samples are checked further for complete compliance with industry requirements such as aging, solubility, adhesion, etc., at the start of each production run and also at a designated time on a weekly basis.

In this discussion we have elaborated only on that phase of the inspection activity in the manufacturing process which deals with statistical quality control. In addition to obtaining samples for the statistical control of the product our process inspectors are constantly

circulating around the production department and are continually checking the product as it is being produced. Our regular inspection organization consists of Process Inspectors, who have the responsibility of checking the product during manufacture, and Finished Goods Inspectors, who are located at strategic shipping points and are charged with the responsibility of checking the product immediately prior to shipment to a customer or to another division of the Roebling Corporation. These Process and Finished Goods Inspectors, paid on a weekly salary basis and represented by a white collar local of the United Steelworkers of America - CIO, report to their respective Divisional Head Inspector, who is a management representative under the direction of the Chief Inspector. The entire Inspection Department reports through the Chief Inspector to the Vice President in charge of Engineering and is thus completely divorced from the Production Division.

In connection with the production and end use of magnet wire and related products we have also used quality control techniques to good advantage in making an extensive comparative study of the product supplied by the various manufacturers of magnet wire to the general trade. This study covered a wide range of sizes in both enamel coated and varnished magnet wire as produced by many different manufacturers. The studies showed clearly the quality level being maintained by the various suppliers and indicated that the larger manufacturers were all conversant with the high quality requirements of the Magnet Wire user. Studies were made on bare wire diameter and overall diameter, dielectric strength, continuity and softness. The studies were very interesting in that they showed clearly that all major manufacturers were attempting to hold their product on the low side of the size tolerance and that all (with varying degrees of success, to be sure) were attempting to exceed industry specification requirements as regards to dielectric strength and continuity. The study on softness was particularly interesting in that it verified, to a great extent, the general comments made by many end users of magnet wire regarding the softness of the wire produced by various suppliers.

Quality Control techniques have been a big help to us not only in holding our quality of product at a high level, but have also improved our relationship with many of our customers.

Let me share with you an experience we had a short time ago which proved the value of our quality control program and improved, to a great extent, our relations with our customer. One particular morning we received a phone call, from one of our larger users of magnet wire, advising us of a serious problem that had been referred to them by one of their customers. It seemed that certain motors were shorting-out in service and indications at the time pointed to the use of defective magnet wire as the cause of the trouble. The trouble had not been traced to the use of our particular magnet wire but our customer wanted us to be alert to the situation. At the time the phone call was received there were several large lots of material in our Shipping Department scheduled to be shipped that afternoon to this same customer. The questions uppermost in our thought were: "Shall we ship?" or, "Shall we hold and reinspect?" We reviewed our quality control records and charts for the period during which the material in question had been made and found it to be equal to our usual high standard. On the basis of this we released that material for shipment with no further inspection. We also visited our customer, advised him

of our decision, and showed him the data and the control charts on which we had based our decision. We further advised that while we did not believe our material to be the cause of their trouble we would be very willing to work with them in their plant or in our own, on their processing or on our own material, until the cause of the trouble had been located and eliminated. The customer was able to locate the cause of the trouble and eliminate it promptly and our wire was in no way involved in the problem. This is just one more proof of the ability of simple quality control techniques to save money, to maintain quality standards, and to cement human relations. Under such control, decisions which have to be made are based on proven facts and not on mere opinion.

The use of Quality Control Techniques has done much to help us in the manufacture of a high quality magnet wire. This statement is not based on mere opinion, it can be substantiated by the following basic facts:

1. It has raised the quality level of the product we are manufacturing.

2. It has reduced the number of complaints received from customers to an extremely low number.

3. It has reduced the cost of complaints per sales dollar to an extremely low figure.

4. It has improved supplier—customer relations.

5. It has saved our company money.

In short, we believe that the use of simple Quality Control Techniques, as we have applied them to the manufacture of magnet wire, is as necessary to the production of a good product as good raw material, good machines and good operators.

The three M's - good material, good machines and good manpower—plus a good Quality Control program are necessary for the efficient operation of any process at a high quality level.

QUALITY CONTROL IN THE MANUFACTURE OF AUTOMOTIVE ELECTRICAL SYSTEMS

W. Frank Lynn
Ford Motor Company

The electric power plant beneath the hood of the modern automobile is a remarkable piece of equipment. It produces a steady supply of electric energy, delivers it where it is needed, and stores it for future use.

Without this power plant, in fact, the modern car would not be possible. We would still be cranking the engine by hand. Lights would be produced by oil or gas lamps. Radios, heaters, air conditioning, power steering and the many other useful accessories which add to motoring pleasure, comfort and safety would be unheard of. Modern transportation, with its tremendous impact on our national economy, would still be a far-off dream.

Perhaps it is a little presumptuous to give so much credit to the electrical system in the progress the automobile has made, but the fact remains that the electrical system is the nerve center without which nothing else will function.

We who are in the business of manufacturing the parts and assemblies which go to make up this nerve center thus carry a big responsibility. These parts must be dependable. They must stand years of hard use — even abuse and neglect — without letting our customer down. That means quality must be closely controlled from beginning to end.

Over the years that Quality Control has been earning its place in American industry, the control of quality in the manufacturing processes has usually received the lion's share of attention — and perhaps rightly so. But in our earnest attempts to keep our manufacturing errors from going out the door, through rigid process and end item inspection, we may have overlooked something very important.

We've all heard it said, time and again, "Quality can't be inspected into a product — it must be built in!"

That's only part of the story. Quality can't be built into a product unless it has been designed in. That's the point some of us have overlooked. That's where quality must begin — on the designer's drafting board. And our responsibility for it cannot end until the finished product is in the customers' hands.

Some of you may feel that design matters have little to do with our function in Quality Control — that we should leave that to the engineer. I feel very strongly that we have a vital interest in proper design, and that we should make our voices heard at every opportunity. We must get across to our designers the interdependence that exists between their job and ours. The finest and most brilliantly conceived design is useless until translated into physical reality. We cannot make that translation efficiently unless the designer has taken into full consideration the limitations imposed upon us by manufacturing processes and by competition.

Actually, for the most part we have a good working relationship with our engineers. The tie between designed quality and manufactured quality

seems to be growing a little stronger every year. You might say that the engagement has been announced, but the marriage has not yet been consummated. So it is a little early to tell just what issue will come from this union.

I want to explore this subject of designed quality more fully a little later on; but now let's go to the other side of the picture, to the job of building the quality that the engineers have designed into our products.

In the Accessory Division of the Ford Motor Company we manufacture the principal equipment which goes into the electrical systems of our automotive products (One major exception is the battery.). In this Division we make the bulk of this equipment, including the voltage regulator, generator, starter motor, ignition coil, distributor and horn. In this discussion we will therefore concentrate on some of the quality control techniques applied at our Ypsilanti Plant.

We are vitally concerned, of course, with a close control of quality at the receiving docks. But our receiving inspection techniques are probably not appreciably different from those in other industries, and there is little point in dwelling on them at any length. We have departed from the Military Standard sampling plans in some cases, and have developed our own Ford plan, which requires a smaller sample number than the Military Standard. We feel, and results bear us out, that it gives us adequate protection for our requirements, at a lower manpower cost.

We keep the normal receiving records for each vendor so that we can tell at a glance how each one stacks up quality-wise over any given period of time.

Our floor and in-process inspection at Ypsilanti is limited to those items which the production operator cannot check, or which would take too much of his time. Where it is possible, the production operator is given the necessary tools or gages to do the job, and time for checking is allowed in the work standard for the operation. This puts a major responsibility for quality parts where it belongs — on the operator. Spot checks are taken by floor inspectors on in-process items according to predetermined cycles and sizes. These checks are taken often enough to give assurance that assemblies will go together. The sample taken for a spot check may vary between five and fifty pieces, depending upon the part or assembly concerned.

This dependence upon production checking for so much of the actual in-process inspection may seem to indicate that Quality Control has abdicated some of its obligation and responsibility. Actually, we find the reverse is true. Production checking tends to create a greater quality-consciousness on the part of production operators, and it reserves to Quality Control the analytical and statistical functions, where we feel that its efforts should be concentrated.

The end item sampling at Ypsilanti is concerned primarily with function and performance, and in some cases with appearance. I will enlarge on some of the phases of this important part of our program later.

All other operations within the Quality Control Department are aimed at helping the shop inspector. The Statistical Analysis Section develops Inspection Instruction Sheets for each part or assembly so that there

will be a consistent inspection pattern to follow. Data sheets which are filled out by the inspectors are returned to the statistical section for analysis of the trouble spots.

The Gage Section handles inspection of initial samples of purchased parts, and checks machine setups to make sure they are correct before a production run is started.

There are certain checks pertaining to the life and performance of electrical equipment that take time to complete, and thus can not be a part of the normal end item inspection. This duty falls to the Physical Laboratory. The laboratory also analyzes field returns. By means of these various tests and analyses, the laboratory may develop information concerning field performance that will be helpful in future production.

Despite all the in-process controls we may wish to install, it is an inherent trait of the mass production system that it will invariably have short periods of producing sub-standard quality, and some of that production may be passed on to the customer. These sporadic periods can spell trouble unless you are governed by an adequate end-item sampling technique. This is particularly true in the manufacture of electrical automotive equipment. A good deal of our Quality Control effort in our Ypsilanti Plant as well as throughout our Division has thus been concentrated on our end-item sampling program in order that we may successfully fulfill our obligation to our customers.

With the aid of several slides, I would like to describe how the general system functions.

Slide 1A - Introduction to Our Reporting System

These forms represent the backbone of our auditing system. The first, the Inspection Instruction Sheet, establishes the acceptable quality level (A.Q.L.) and outlines the inspector's duties. The second, the Quality Control Work Sheet, affords him a rapid method of recording his data. The third, the Outgoing Quality Trend Chart, is a simple and graphic method of reporting and evaluating the quality level to management and to the customer.

Let's look at each of these forms individually.

Slide 2 - Inspection Instruction Sheet (Form 732)

This form is directed at the inspector. It supplies him with the necessary information required to operate his station by listing what he must check and with what equipment. It lists the characteristics and the gage requirements as well as the quality level and frequency of inspection.

The plant Quality Control analyst originates this form. He evaluates the product and establishes the quality standard which must be maintained to produce an acceptable end item. We attempt to do this in conjunction with the customer prior to production. In our case, the customer is one of our own Company Divisions. However, more often than not, we must originate the entire form and establish the quality level, and then at a later date arrive at a mutual agreement as to characteristics to be sampled and quality level to be maintained.

In general, this last method is not too great a handicap. We have found over a period of time that the discrepancies, if any, are minor in

Slide 1A - Introduction to Our Reporting System

QUAL. CONT MGT PEM-732

PARTS AND EQUIPMENT MANUFACTURING DIVISION

INSPECTION INSTRUCTIONS

PLANT Ypsilanti #1 PAGE 1 OF 1

PART NAME	LATEST E/P DATE	DATE	PART NUMBER
Generator Assemblies		1-2-57	All 10000

RECEIVING ☐ IN-PROCESS ☐ END ITEM ☒

END ITEM IN WHICH PART IS CONTAINED Same

SAMPLING PLAN USED	LOT SIZE	SAMPLE SIZE	ACCEPT. NO.	REJECTION NO.
MAJOR AQL C		30	2	3
MINOR AQL D		30	3	4

NO.	CLASS.	CHARACTERISTIC	INSPECTION METHOD	GAGE NO.
1	C	Assy must be free of missing or incomplete operations	Visual	
2	C	Assy must not be damaged, mutilated, etc.	Visual	
3	C	Assy must have required output	Allen Tester	
4	C	Assy must not be noisy	Allen Tester	
5	C	Assy must not be tight	Allen Tester	
6	C	Distance between bosses	Length Bar Gage	24ZF-1068 24ZF-2384
7	C	Oiler and cap must be present (6 cylinder)	Visual	
8	D	Terminals must not be damaged	Visual	
9	D	Mounting boss and terminals must be free of paint	Visual	
10	D	End plate must be down	Visual	
11	D	Assy must have proper pulley	Visual	
12	D	Pole screw torque 175-200 in. lbs.	Torque Wrench	
13	D	Thru bolt torque 55-75 in. lbs.	Torque Wrench	
14	D	Terminal screw torque 15-35 in. lbs.	Torque Wrench	
15	D	Assy must have proper and legible name plate	Visual	
16	D	Assy must have proper decal	Visual	
17	D	Assy must be painted properly and free of sags or chips	Visual	

ALL SPECIFICATIONS ARE SUBJECT TO PERIODIC LAYOUT INSPECTION AND/OR LABORATORY ANALYSIS

103

Slide 2 - Inspection Instruction Sheet (Form 732)

nature and are usually a matter of interpreting a specification which may not be clear-cut; for instance, how much burr may be allowed, what is flat, how sharp is a sharp corner, and so forth.

Slide 3 - Work Sheet (Form PEM-7337)
After furnishing the inspector with a method of inspection, we then give him a rapid system of recording the data. The numbers on this sheet correspond to the characteristics on the Form 732 so that a quick dash mark denotes the occurrence of a defect. Through a simple tabulation he may quickly render a decision as to accepting or rejecting the lot.

At the end of each work shift these sheets are given to the Quality Control analyst assigned to the product. From these he may calculate the process average of the end item, or of each individual characteristic sampled.

Slide 4 - Outgoing Quality Trend Chart
This represents the culmination of our reporting system. It is a simple graphic presentation of the history of quality by which persons on management level may interpret quality trends of the individual products.

At the end of each work week the analyst computes the data as recorded on the work sheet. It is then duplicated and copies are sent to all management personnel directly concerned with the quality of the particular product.

Summary of Slides 2, 3 and 4
Once again, let us briefly review.

Step #1 - Instruction Sheet, sets the quality level and spells out the inspector's duties.

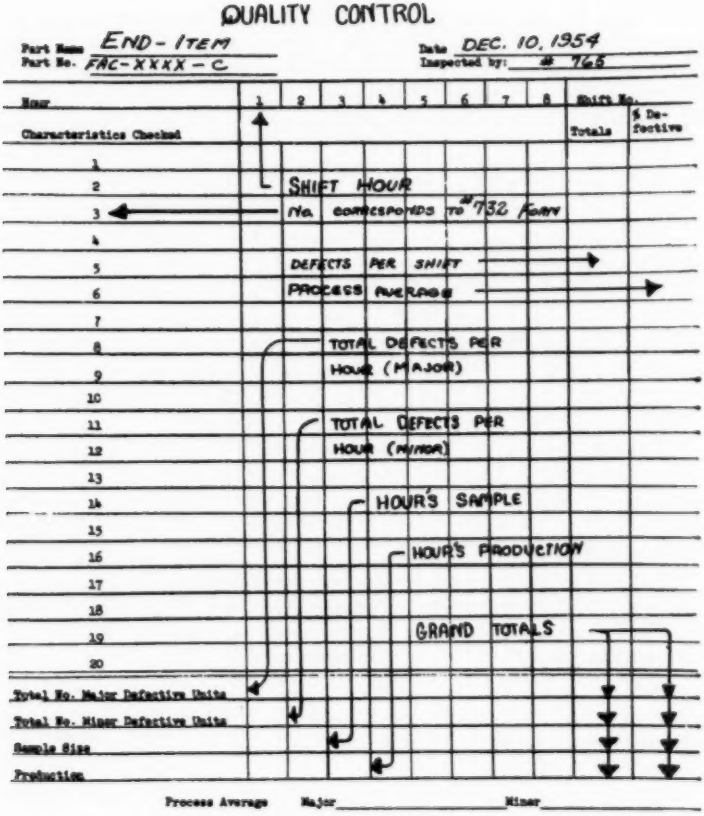
Step #2 - Work Sheet, provides a method for the inspector to record the data.

Step #3 - Trend Chart, provides a reporting and evaluating method for Management and, in some cases, the customer.

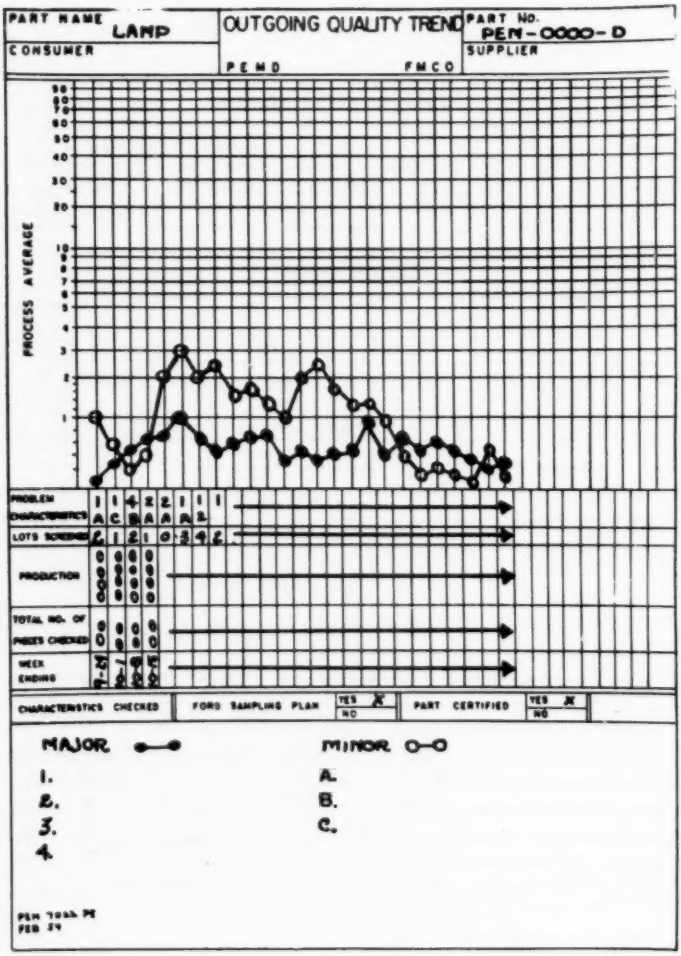
This same system, with slight modifications, is used in our receiving inspection area.

In our attempt to develop a standard end-item sampling procedure to which the entire Division could conform, we found that we could not set one pattern for the actual inspection method. For example, let's take the generator, which you might say is the heart of your car's electrical system.

Inspection Instructions for the generator assembly list seventeen different characteristics that must be checked by the end item inspector. Of these, the most vital characteristic is the electrical output. This is of such primary importance that it must be checked 100% in process. This is done by production fixtures which merely indicate by green and red lights that the unit is either satisfactory or below specification. If satisfactory, it is placed on a chain conveyor, passed through the paint booth, and then sent to the shipping area. There the end-item sampling takes place. The units are sample inspected by a variable tester, at a test stand, and then loaded for transport.



Slide 3 - Work Sheet (Form PEM-7337)



Slide 4 - Outgoing Quality Trend Chart

Unfortunately, some lots did not pass this sampling inspection and the checking capacity of the test stand was only a fraction of the total production. It becomes necessary to find a suitable method of screening these rejected lots. I would like to emphasize at this point that the Production Department must be charged with this screening operation to make the program effective. This creates a burden and penalty and, as such, serves as an incentive to correct the characteristics causing the rejection.

Another place where we have a 100% checking operation, plus end item sampling, is the voltage regulator. The regulator, as you know, controls the amount of current the generator produces, and lets it make just enough current to meet the operating requirements of the electrical system. Without the regulator, the battery and other electrical units would soon be permanently damaged, and would, of course, cease to operate.

If the generator is the heart of your car's electrical system, then certainly the voltage regulator is the brain; it must operate in temperatures that may range from Death Valley's 120 degrees to 19 below zero or lower, such as many states experienced this winter. It must operate under varying power loads, from the bare essentials of electric equipment found on the standard models to those deluxe jobs with power steering, power seats, power window lifts, air conditioning, and so on. The regulator has to be calibrated within a range that will serve either type of car, so that just the right amount of current will be available in either case. That means the blue print specifications must be rigidly adhered to, or in a short time we are going to have some mighty unhappy customers.

The regulators are calibrated "hot"; that is, under conditions which simulate normal running temperatures. Under these same running temperatures they are then 100% tested by production checkers, visually and electrically. Quality Control then takes its end item sample, consisting of twenty-four (24) pieces every half hour. These are put on a series of monitor stands, hooked up with a generator of the specified capacity and are checked for thirty (30) minutes at normal running temperatures. This gives the regulator the same kind of test it would get if placed on a car for that length of time. If the regulator is not properly calibrated, it will show up in this test. Checks are made for output, reverse current and cut in, and for temperature compensation. Lots are time stamped for identification, so that the full thirty (30) minute test is assured.

The same pattern is followed as for the generators: satisfactory lots are loaded for shipment, and rejected lots go back to production for screening.

Our end item sampling often helps us straighten out problems encountered in previous operations. Some time ago, in making a frequency distribution of the results of our regulator end item check, we found we were getting a non-normal distribution. We went back to the 100% production check and broke down the check according to individual operators, identifying each one through a color code. We found that the biggest share of end item rejects were coming from one certain operator of the four who run this check. A second operator showed almost no rejects chargeable to her. The other two had about the normal amount to be expected.

We found that Operator A seldom rejected a regulator at her station, while Operator B was so extremely critical that she was rejecting

practically everything. A thorough discussion and explanation of the results we were after served to clear the matter up with Operator B. As for Operator A, further investigation showed that she actually could not see well enough to read the gage properly. She was transferred to other work where sharpness of vision was not so essential.

We used a similar study to determine the effectiveness of the operators doing the calibrating on these regulators. A 15-piece hourly sample was taken from the 100% inspection station for each calibrator. Analysis of the information showed that one of the nine calibrators checked was having between 5% and 12% of her work rejected per day — an average of 8% for a three-day period. The other eight calibrators had from 2% to 3% rejected.

The hourly sample indicated that this operator was adjusting from the high to the low limits — that is, if she found that the regulators she had calibrated for the past hour were near the low limit, she would work toward the high limit. She reviewed her work at the 100% inspection station more frequently than the other operators in an attempt to stay within the limits. Consequently, she was adjusting more often and, therefore, was subject to more adjusting errors. The wider distribution she was getting was therefore not because of poor workmanship, but because of too much changing of her setting point to prevent poor calibrating. This is one of those rare cases where an over-conscientious person is breeding a condition she is actually trying to prevent.

The final solution in this particular case was to transfer this operator to a different job where her extra-conscientiousness was an asset rather than a handicap.

Because of the extremely critical role it plays in the automotive electrical system, the distributor also comes in for a rigid 100% check by Production, followed by end-item sampling inspection to assure a quality product for the customer. The 100% check is performed simultaneously with the calibration of the distributor by means of an ingenious test machine. In this case, we get two operations for the cost of one. This machine, a cathode ray oscilloscope, checks electronically for point gap setting, mechanical advance, and vacuum advance. Calibration adjustments are made by the operator in accordance with the results registered on the screen of the machine.

The electronic tester is a great advance in the inspection of distributors. Before, we were obliged to make a point-by-point check of distributor performance, which was at best a laborious and time-consuming process. Now, we are able to see at a glance the whole performance pattern and make our adjustments accordingly. The operator, in reading the screen, can detect such assembly defects as bent shafts, bad cams, bad breaker plates, and faulty point assembly. In such cases, the operator notes the defect on the assembly by means of a code number, and the unit is sent to the repair bench for correction.

After end-item sampling, satisfactory lots are shipped, and rejected lots are returned to Production for re-checking.

The importance of the distributor in the satisfactory performance of the automobile imposes the necessity of holding to a 1% Acceptable Quality Level on important characteristics. There are fourteen (14) major characteristics that are cause for rejection at the 100% production check

or at the end item inspection, and you can appreciate the way the problem of inspection is intensified all the way along the line in order to come up with a satisfactory A.Q.L. at the end. Undersize cam holes, bad cam contours, burrs on cam plates, cams binding on shafts, shorted point assemblies, bad springs, cracked retainers, defective diaphragm assemblies -- all these require strict in-process controls in order to avoid teardown and repair of the completed units at the end of the line. Here, again, is why we stress the role of the production worker in the quality program, depending on him to keep his own operation properly policed at all times, with the floor inspector spot checking at stated intervals to maintain sampling control.

In our ignition coil and horn departments we follow a similar pattern of control through the line, with 100% electrical testing by Production, followed by end item sampling by Quality Control.

To sum up our philosophy, if you want to call it that, of Quality Control for our Division, we use the techniques that have proved themselves over years of application, but we do not hesitate to depart from those techniques to meet new situations. We have tried to reconcile the practical and economical phase with the theoretical aspects of the field. We have attempted to develop a framework broad enough to satisfy the diversified nature of our Division.

Now I would like to go back for a few minutes and take another look at this business of designing quality into our parts ... the other half of this marriage we hope will take place one of these days.

Let me say right now that the designer has his problems, whether he is designing voltage regulators, motors or widgets. His aim is to control quality, performance and life of a product by maintaining correct fit, alignment and clearance of parts. He wants to make sure parts and assemblies are interchangeable, and to provide a factor of safety against the possible use of parts outside the specification limits.

On the other hand, he knows that it costs money to hold to close tolerances. It costs more for pre-production planning, tool design and tool making. It costs more for tool and machine set-up, maintenance and replacement; more for gages and inspection time. It creates more scrap and requires more rework. It requires more operations and greater training of personnel.

Designers are coming more and more to realize the importance of arriving at a proper compromise in the matter of tolerances, so that the producer of the part can make it in sufficient quantity, and at a competitive price. They realize that where dimensions directly affect the performance of the product, the main criterion should be the functional requirements. Conversely, where dimensions do not materially affect the function, the main criterion should be the production requirements.

The roots of the tolerance problem go deeper than just making the production or assembly job easier. What we as Quality Control people need is not simply more realistic tolerances, but a more faithful observance of those tolerances. It isn't a matter of quality versus quantity, so much as it is getting the most of both by means of more efficient quality standards. We can't get away from the fact that every time a tolerance is found by production or inspection to be unnecessarily close

on a part where proper function is not affected, there is a lessened respect for all other tolerances specified. On the other hand, every time production or inspection finds that they run into difficulty because a drawing tolerance is not followed, the effectiveness of the other drawing tolerances is increased.

I think you will agree that a number of things could be done to foster and promote a better mutual understanding of the problems of both the designers and the production people.

Designers should get together and define the relationship between manufacturing variations and product performance, and agree on the ideal fit for a given function. They should then present this basic data in such a way that their colleagues in the plant can better appreciate the designer's point of view.

Production engineers should get together and define the relationship between tolerance and production efficiency, and present to the designers reliable information as to just what is practical and economic from the production standpoint.

At this point, Quality Control can aid immeasurably by making available to Engineering such techniques as process capability studies, root-mean-square concept of tolerancing, and statistical techniques of applying A.Q.L.'s by mating part fits.

Then let Management oil the wheels of collaboration by lending to the task the weight of their authority and encouragement, and by exercising their function of coordinating opposing interests.

This is a job where everybody must visualize the advantages of the end result — and work together to attain it. The objective is worth while, gentlemen. I invite your active support in bringing the importance of designed quality into its true position alongside manufactured quality.

Let's hope that marriage date isn't too far away!

DEVELOPING QUALITY STANDARDS AND RATINGS IN FOLDING PAPER CARTON MANUFACTURE

Robert B. Benson
Continental Folding Paper Box Co., Inc.

Introduction

During the past few years, important Quality Control advances have been made in the manufacture of Folding Paper Cartons. These advances have resulted from applications of modern Quality Control methods with modifications to meet requirements in our industry.

For accurate evaluation of folding box quality, measurable quality standards are needed both for appearance and for functioning. I would like to describe in this paper the methods employed at Continental Folding Paper Box Company in establishing and maintaining such measurable standards for our products.

As stated above, quality characteristics must be controlled in the following two areas:

A. The Appearance: In our mid-twentieth century self-service merchandising establishments, the package has become a salesman for the product. Each package must stand side by side with numerous competitors and virtually say to the customer "Take me!" As such the carton must consistently maintain in "high fidelity", if you will, all of the original intent of advertising and packaging promotion designs. This requires close control of color, register, half tones and general appearance.

B. The Function: In addition to containing the product safely through handling and transit, the carton must have the feature of machineability which will enable it to be filled on modern high speed packaging machines.

The quality conformance in these two areas is judged largely in terms of attributes. Quality Control therefore requires the creation of visual standards and functional tests.

Defect Classification

Our Quality Control system is based upon a defect classification system for the evaluation of both the appearance and functional defects. This classification follows the procedures in U. S. Department of Defense Sampling Plans.

A "minor defect", for which we penalize one demerit, is one that does not materially reduce the usability of the unit of product for its intended purpose, or is a departure from established standards having no significant bearing on the effective use or operation of the product.

A "serious defect", for which we charge five demerits, is a defect that could result in failure, or materially reduce the usability of the unit or product for its intended purpose.

We define a "critical defect" (penalized 10 demerits) as one which would cause a customer substantial trouble in filling the carton, or in the use of his product by the customer.

CHART # 1

	Date													
	4/18	4/19	4/20	4/21	4/22	4/24	4/25	4/26	4/28	4/29	$\sum$	$\bar{c}$	$w \cdot \bar{c}$	$w^2 \cdot \bar{c}$
Number of Critical Defects	0	0	0	0	0	0	0	0	1	0	1	0.1	1.0	10.0
Number of Serious Defects	1	1	0	0	3	1	0	0	0	0	6	0.6	3.0	15.0
Number of Minor Defects	8	19	9	4	8	7	7	6	7	7	82	8.2	8.2	8.2
Sample Size	1	1	1	1	1	1	1	1	1	1	10			
												Σ	12.2	33.2

DU = Demerits per Unit

$$\overline{DU} = \Sigma(w \cdot \bar{c}) = 12.2 \text{ or } 12$$

$$\sigma_{DU} = \sqrt{\Sigma w^2 \cdot \bar{c}} = \sqrt{33.2} = 5.8$$

$$UCL = DU + 3\sigma_{DU}$$
$$= 12.2 + 3(5.8)$$
$$= 29.6 \text{ or } 30$$

Reference Quality Samples

The basis for classifying visual defects is a standard book containing sets of samples for each defect. Every set of samples has one carton showing the degree of defect for each severity rating. Therefore our minor, serious, or critical defects are defined by example in a reference book available to Inspection, Production and Sales personnel.

I think it is an axiom of Quality Control that much more of our time is spent in defining tolerances and standards pertaining to visual defects rather than measurable defects. It is harder to argue with the results of a micrometer, voltmeter or tensile test than to disagree with a visual evaluation. By defining our standards with reference samples, much disagreement in this area has been avoided. The defect classification books are applied to all orders from cake boxes to biochemical cartons. Acceptance for our different customers is based on the number of minor, serious and critical defects each customer specifies as acceptable. In other words our standards are uniform but the specifications vary for each job.

Control Charts for Processing

In order to maintain these quality standards throughout our operations, we have established process inspection in all manufacturing departments. Here the product is continually checked against the same quality standards which will be used to judge the finished product. We have been able to prevent defectives from being produced and have steadily improved quality in manufacturing areas by this means.

Standard control charts are maintained at each of our presses for the information of operation and manufacturing supervision. These are "C" charts which rate quality as the number of defects per sample. In our case, they apply to the number of defects per sheet of paperboard. Since we are dealing with three levels (or classifications) of defect severity, an adaptation of the "C" chart originated by the Bell Telephone System is used; this is described by Mr. H. F. Dodge in his article on "A Method of Rating Manufactured Product" published in the Bell System Technical Journal April 1928. The computation for Demerits Per Unit and the Upper Control Limit are shown in Chart #1.

On this chart we have tabulated the total demerits under each severity classification using average demerits for one day as one sample. These represent approximately sixteen inspection visits per day.

The last four columns represent the calculations necessary for the computations below. C bar is the total defects divided by the sample size, in this case, ten. W times C bar is the product of the defect severity rating and the average defects; this gives the average demerits for the class. The last column shows the product of the severity squared, times the average defects for each class.

The Demerits per unit or sheet than is the sum of the demerits for the different defect classes.

The Upper Control Limit is calculated by finding the standard deviation from the last column on Chart #1. The UCL then is equal to av. dem-unit plus 3 times the standard deviation. In the example shown here the average demerits per sheet were 12 and the upper control limit was 30.

CHART #2

CONTINENTAL PRINTING MACHINE RECORD SHEET

ORDER NO. 9324 CUSTOMER ——— JOB DESCRIPTION 3 Colors 12 Cartons 18.8 #'s PRESS NO. 4

PRESSMEN Jones Smith DATE: 2-16-56

COUNTER READING

TIME: 8:00 9:00 10:00 11:00 12:00 1:00 2:00 3:00 4:00 5:00 6:00 7:00 8:00 9:00 10:00 11:00 12:00 1:00 2:00 3:00 4:00 5:00 6:00 7:00

OFFSET

Time	Total Cartons Chkd.	Total Cartons Def.	% Def.

— TOTAL DEMERITS —

36 32 28 24 20 16 12 8 4 0

UCL

DU

Defects																								
Plate Batter & Cats Eyes	5																							
Weak or Hard Impression			10																					5
Dead Metal									5															
Mottle																								
Register																								
Worn Plate																								
Fill-up Reverse Type																								
Fill-up Screen			5																					
Wax																								
Spots																								
Smears																								
Off Color			5																					
Other																								
Total	8	11	32		7		11	15	18	6	12	16		11	4	10	17							

Chart #2 shows the application of this technique to the control chart. After some experience with this chart we became acutely aware of two shortcomings:

1) Data had to be collected over substantial periods of time before statistical control could be established. In a job shop such as ours, this is a serious handicap since many orders may run for only a few days.

2) The control limits could apply to only one job or set of specifications.

Pre-Setting Quality Levels

To overcome these shortcomings, we have used our experience with past runs to develop a method of pre-setting quality levels for future production on new jobs. This has been done by establishing a standard formula for such quality calculations.

By means of multiple correlation techniques, using the essential variables which influence product quality, we have established this formula and can now set up immediate statistical control. The calculations for this prediction are simple algebriac solutions performed with ease by our inspectors.

Since our production operations are divided into three basic areas: Printing, Cutting and Creasing, and Finishing, we maintain an inspector in each department covering both shifts. The modified "C" chart with pre-set limits is utilized in the first two sections.

Quality Audit

The Finishing operations are evaluated by an hourly Quality Audit which can be seen in Chart #3. All of the possible defects which may have occurred throughout the plant are covered. We are able to compute an average demerit per carton figure and a percent defective estimate.

We have control by die position number, an important feature for isolating defective material. Furthermore we can determine what percent each defect is contributing to the quality level. These sheets are summarized weekly and monthly for the purpose of feedback to Production Supervision and Management. The measures exercised to maintain our dual requirements of appearance and function extend into other areas than process control.

Laboratory Control

With full realization that quality workmanship cannot be realized when inferior materials are used we have established a full time laboratory control of vendor goods.

Incoming paperboard shipments are randomly sampled for moisture, weight, caliper (or thickness), stiffness, water absorption and other important characteristics. The results are recorded on a laboratory worksheet and analyzed for conformance to standards and specifications. Shipments which do not meet requirements are reported to management with our recommendations for disposition. At the same time the defective material is tagged and held until such time as that disposition has been finalized by management.

CHART #3

ORDER NO. 9324 CUSTOMER

CONTINENTAL OUTGOING QUALITY AUDIT

STYLE NO. 1110-3 JOB

Summary Sheet

M/C No. 8
Inspector: G.P.
Date: 11-30-66

APPEARANCE DEFECTS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	28	29	30	TOTAL	%
Plate Better																									10	1.7
Color Variation																									25	4.4
Cat Eyes																									44	7.7
Impression																									30	5.2
Smears																										
Dead Metal																									12	2.1
Mottle																									1	0.2
Offset																									41	7.1
Print. Register																									98	17.1
Cut. Register																									5	0.9
Screen Fill Up																										
Wax Spray																									121	21.1
Belt Streaks																									11	1.9
Print. Streaks																									2	0.3
Trapping																									2	0.3
Reverse Type																									10	1.7
Plate Wear																										
Other Bent Flap																									40	7.0
FUNCTIONAL DEFECTS																										
Blood Dimension																									40	7.0
Dry Glue																									5	0.9
Glueing Off Line																									25	4.4
Incomplete Glueing																										
Poor Perforation																										
Cracked Score																									1	0.2
Punched Score																									20	3.6
Stripping Damage																									30	5.2
Feeder Point Scuff																										
Tear																									1	0.2
Tympan Crush																										
Wobbed																										
Poor Cutting																										
Poor Scoring																										
Other																										
Total	70	109	104	89	80																				574	
Number Cartons Insp.	29	21	28	25	32	2																			160	
Average	2.5	5.2	3.7	3.6	2.5	1																			3.6	
Number Rejectable					1																				1	
% Rejectable					3.1%																				0.6%	

Paperboard suppliers provide us with a copy of their own quality control data sheets together with control charts so that their testing can be verified against our own results.

The two sets of data provide us with information suitable for establishing process capabilities and for product improvement using the statistical methods of Quality Control.

Controls over inks and other materials are exercised in a similar manner. Conformance with color standards and resistance to fade, friction and packaged product are particularly important characteristics of ink which must be maintained.

Quality Control Engineering

At the same time we engage in as much Quality Engineering as time and opportunity will permit. Under this division we evaluate and contribute to process improvement, experimental investigations, the establishment of new quality standards and new product quality.

Results Obtained

Of the many benefits realized as the result of our program, the elimination of continuous 100% inspection on key accounts at the Finishing Department was perhaps the most impressive. Our scrap and rework figures were substantially reduced as were the overall quality costs. The best measure of our success lies in our customers satisfaction, which I am pleased to report reflects itself in their favorable comment and continuously increasing order placements.

PROBLEMS AND PITFALLS IN ACCEPTANCE SAMPLING OF GLASS CONTAINERS

K. F. Lang
H. J. Heinz Company

This paper will discuss some of the practical problems that must be recognized and considered in setting up an acceptance sampling procedure for glass containers. The discussion will be confined primarily to acceptance sampling based on inspection by attributes. To examine every individual container in a shipment, or run select tests on every case of containers, would require a laboratory force out of proportion to the risk involved. It is here that a well defined statistical sampling program can be applied at a minimum inspection cost to assure against acceptance of unsatisfactory glass containers. Any proposed plan of sampling inspection, no matter how well defined, requires willingness to take a chance. Granted, statistical control techniques can limit such risks to some degree.

Some glass companies, while not opposed to user statistical sampling programs, are reluctant to promote the adoption of such programs by their customers. They are inclined to think of their own procedures as being more than adequate to assure their customers good "commercial quality". Customer sampling programs in their opinion, results in duplication and added costs for which there is little or no justification.

On the other hand some users of glass containers feel they have the answer to all their problems by the adoption of an acceptance sampling program. Even the best is not a panacea.

Through cooperative efforts however, a mutually satisfactory program can be established. This will require frequent meetings with glass suppliers, particularly during the initial stages, so as to prevent misunderstandings.

With these thoughts in mind, I would like to discuss with you today, some of the problems and pitfalls in acceptance sampling of glass containers. Our remarks will be based on experiences we have had with acceptance sampling of glass containers. We do not intend to explore the mathematical aspects of acceptance sampling but hope to convey some of the practical problems that have come about through experience.

Our present program of acceptance sampling of glass containers was introduced during the early part of 1953. This was preceded by a year of investigational work during which time various glass suppliers were contacted in an effort to consolidate our thinking and come up with a mutually satisfactory program. There were many details that needed clarification. (Slide I) The simple matter of terminology was an early stumbling block. Not only were we confused by terms used by individual glass suppliers, but were surprised to learn of a lack of common terminology within the glass industry. One supplier frequently did not know what another supplier meant in describing a particular type of defect.

By December of 1953 acceptance sampling of one basic type glass container had been introduced to all phases of our business. Although the acceptable quality levels established at that time were considered tentative, they have proven to have been practical with time and experience. We can say with some degree of satisfaction that the general quality

level of glass containers has been improved through such a program. It is further felt that through such improvement has come a reduction in production downtime, loss of product, and consumer complaints.

Management Approval - Selling the Program

One of our first problems (Slide II) was selling the need for acceptance sampling to top management and then selling it down the line. Fortunately, as in most aggressive companies, a sound acceptance sampling program was readily accepted.

Prior to 1953 the receiving inspection of glass containers for use in our various plants, was dependent primarily on the knowledge developed by certain individuals from experiences over a period of years. Thus, the degree of inspection at a particular plant differed from that of another plant because it was dependent upon an individual. Glass containers accepted at one plant were subject to rejection at another plant or vice versa due to a lack of uniformity in methods of sampling, inspection, and/or interpretation of percent defects found.

Such a program for many years seemed to have merit. With an experienced glass inspector the evaluation (so called) of a glass shipment could be arrived at rapidly. Sampling was usually relatively small, thus minimizing the time and costs involved.

On the other side of the ledger we noted a tendency by individual inspectors to sometimes minimize major defects and exaggerate the importance of minor defects. With different inspectors having varying degrees of experience at various plants, there was an obvious difference in acceptance and rejection levels. We also found that hasty judgment fostered by small samplings was unfair to the supplier and at times costly to our company. This procedure likewise gave grounds for debate which led to embarrassment for both supplier and user. An effort to establish a better relationship with our various glass suppliers and to insure uniform quality in the glass containers received, led to the development of an acceptance sampling program based on Mil. Std. 105-A. Various meetings were held with technical representatives of glass suppliers so as to gain the benefit of their experience.

The result was a program readily acceptable to both management and glass supplier. Management was concerned with the initial cost but the improvement in quality and line performance justified the investment.

Putting the Program into Action

The launching of any successful program requires careful preparation. This is certainly true with an acceptance sampling program.

An important step (Slide III) is that of obtaining agreement between the user and supplier of glass containers. Bear in mind, the user is desirous of perfection with resultant good line performance. The glass supplier is conscious of the inspection costs involved to maintain a user's desired quality.

Once an agreement is reached it is necessary to spell out the program so that it can readily be absorbed by all individuals responsible at both the user and supplier's plants.

Improper instructions, lack of definition, inadequate forms, and insufficient supervision can undermine the program right from the start.

Classification and Definition of Defects

Evaluating the relative importance of different types of defects can be a problem. (Slide IV) There must be agreement between the user and supplier as to the seriousness of certain defects. This is not a simple matter when discussing glass containers. For example, size, degree and location of a defect and its resultant effect on possible breakage must be considered. Furthermore, how serious is a body blemish to the final end point of the sale of the merchandise?

As each type of defect is evaluated one must first arrive at a common understanding as to cause and resultant effect on strength of the container. This is not always easy to do.

With each defect arises the necessity for agreement with suppliers on the significance of that particular defect. Many times it has been necessary to make tests on our filling lines with selected defective containers so as to confirm an opinion.

Even after agreement has been reached on a tentative classification the entire listing of defects is subject to revision based on laboratory data and/or line performance data. Some of our present defect classifications have been revised at least five times.

Selection of Limit Samples

Closely allied with the classification of defects (Slide V) is the establishment of "limit samples". There are limitations regarding the ability to define a type defect with words alone. Where possible, a limit sample is used to show visually to what extent a defect can occur and still be acceptable. This also enables both the user and supplier to show inspection personnel the degree of permissible deviation.

Obtaining agreement on "limit samples" can be a problem. The user is concerned with line performance and potential consumer complaints while the supplier is concerned with his ability to sort out ware to the prescribed limit sample at a normal operating cost. Frequently, a closure supplier will test a particular finish defect to determine the efficiency of the seal. Limit samples of this type can definitely be selected on performance. Over the "bartering table" agreement is usually reached at least on a tentative basis. With experience and through usage limit samples can be modified by agreement.

There is also the problem of availability of "limit samples". The user is dependent largely on the supplier for the accumulation of samples. A few can be selected by the user on his filling line inspection. However, the best source is from the suppliers reject ware. Still, this may develop into a long term project before a complete set/or sets of exact limit samples can be accumulated.

One cannot over emphasize the importance of limit samples in the possession of both user and supplier. Many controversies can be avoided through their use.

Establishment of Acceptance Levels

Acceptable Quality Levels (AQL's) must be realistic and acceptable to both user and supplier. (Slide VI) It is sometimes difficult for executives or even production supervisors to understand why one would accept anything but perfect lots. This is aside from the fact that they have been working with much less than perfection and that whether they like it or not something less than perfect lots will always have to be tolerated. As stated by a well known statistician--"You simply cannot afford either to ship to your customers or insist on buying from your suppliers nothing but perfect lots. The price of perfection is simply too great."

There are always two views to be considered in setting up an acceptance level--the view point of the supplier submitting the product for acceptance and that of the user who must decide whether to accept or reject the shipment. The user requires protection against the acceptance of too many defective containers. The supplier, on the other hand, needs to be protected against the rejection of too many good containers. He is also concerned with the use of an acceptable quality level that he can live with economically.

Following repeated meetings with suppliers as well as our own management groups we arrived at tentative AQL's which were submitted to our suppliers for approval. Some suppliers accepted them readily, others with reluctance, and still others on a tentative basis only.

In setting up acceptable quality levels for glass containers one has to consider a number of variables which have a definite bearing on the ability of the glass supplier to produce ware relatively free of defects. We now enjoy a rather high degree of acceptable quality in certain containers primarily because they are manufactured on a year-round basis. This enables glass suppliers to utilize one machine continuously with trained operators. It also permits them to concentrate their efforts on improving methods of manufacturing and inspecting of finished containers. Another factor is that of design. A plain round design glass container lends itself to a better quality glass container.

One must not forget the ultimate use of the container. This has a definite bearing on the demands of the user for AQL's sometimes considered unreasonable by the supplier.

Sample Size

Statistical sampling gives us a means of determining the quality level of glass containers by inspecting a relatively small portion of the lot. Unfortunately, emphasis too often is placed on the "relatively small" sample. (Slide VII)

Many users of glass containers, especially the smaller ones, cannot afford and do not wish to spend any more time than necessary on receiving inspection. The cost is too high. Therefore, they are always looking for some short cut which generally is at variance with the proper statistical principles which apply to sampling. There is no known procedure for short cutting sound statistical procedures without markedly reducing the efficiency level of the procedure.

Sample Selection

More important (Slide VIII) than the size of the sample is how and where the sample is selected. This can be a major problem or pitfall in the whole acceptance sampling program.

Inaccessibility of portions of a shipment usually results in a sampling somewhat less than truly random. Sampling practices which deviate from truly random will cause a misrepresentation of the shipment, either favorably or unfavorably.

When a shipment of glass containers is rejected the supplier invariably will mention differences in evaluating the samples and/or non-random sample selection.

In the selection of the sample of glass containers the inspector often faces definite physical difficulty, particularly if a decision must be rendered before the shipment is unloaded. Some method must be developed in selecting the sample without bias from all parts of the shipment. True, physical difficulties must be overcome, at the same time we must guard against inspectors doing it the easiest way. As one authority has stated, "Perhaps our greatest difficulty in making the sample a random sample is human laziness". It is of utmost importance that every effort be made to secure a true random sample since it is fundamental to good acceptance sampling procedures.

Knowing these facts, the user immediately looks for some substitute to avoid the high cost of strict random sampling.

Modification of sampling procedures by way of greatly reduced sample size plus lack of random selection will void the statistical accuracy to a degree that will make the plan worthless.

Glass containers are usually classified as stratified materials. It is therefore unadvisable to take a large group of cases from any particular portion of a car or truck shipment. In order to obtain a fairly good random sample, it is necessary to sample only a few bottles from a case, from many cases selected throughout the shipment.

In order to overcome the cost factor involved in good random sampling we developed a sampling procedure for direct shipments at the time of unloading. It is agreed with the supplier that if the load is rejected he will stand the cost of reloading. Pre-season ware being manufactured and stored for us by the supplier is random case sampled for us. Glass containers are then random selected from such random sample cases at our plant. While sampling and inspecting for visible defects, our inspector sets aside a random bottle or jar for each mold cavity to be used for dimensional inspection.

Physical and Mechanical Aids for Inspection

A problem sometimes develops from inadequate or improper tools as well as physical surroundings. Insufficient lighting may cause an inspector to improperly evaluate visual defects. On the other hand special lighting and/or the use of magnification will be questioned.

In a multiple plant operation it is essential that all plants con-

cerned use similar inspection techniques so that glass containers will receive the same type of inspection at all points. Lack of uniformity in physical and mechanical set up can lead to problems.

Administering the Acceptance Program

The acceptance sampling costs include not only testing and inspection costs but also the costs of administering the acceptance program. One must justify the additional cost of inspection in terms of better quality ware and subsequent better line performance. Quite frequently a problem arises through inadequate supervision of the program. Assuming the correct number of glass containers has been selected in a random manner, they must then be accurately and impartially inspected and the results properly recorded. Correct tabulation of results, feed back of information to the supplier, development of proper forms, can become problems if not administered in the correct manner. Laxity on the part of individuals responsible for the program can seriously weaken its effectiveness.

An effective sampling program demands the quality of supervision that will insure the glass containers being carefully inspected and results accurately recorded.

People to do the Job

Selection of the right people (Slide IX) to do the job can be a problem. They must be exact and patient and have a sense of honesty which enables them to impartially examine each lot. They must not be prone to complacency due to the repetitive nature of sampling inspection. They must be of a nature which avoids forming a biased opinion of a supplier which would influence the evaluation of samples. The best sampling procedure, exact in mathematical detail, is doomed to failure if the right persons are not selected for the job.

One must have confidence in people selected to do the job. We must have assurance that samples will be selected according to a prescribed procedure, examined without bias, and recorded as found. Furthermore, that they must not succumb to human frailties and find "short cuts" or "easier ways to do the job".

One must be certain that persons chosen to do the inspection work are properly prepared and instructed. There are many sources of confusion as the novice looks on acceptance sampling for the first time. There is a vocabulary to be learned. Usually there is some adjustment of thinking in connection with samples and the meaning of samples.

Keeping the Program Up to Date

January, 1957 began the fourth year of attribute acceptance sampling and inspection of glass containers by our company. Since its origin, the program has been expanded to include twelve different types of glass containers.

We are convinced that through such a program we have been able to build up a better supplier-user relationship, have raised the general quality level of glass containers received, and in turn have witnessed remarkable improvement in line performance. The program is in need of

further development to include various other containers used in our business. At the same time there are similar programs in progress covering other packaging materials such as tin containers, metal closures, labels, etc. Although our objectives are clearly defined, there are limitations as to the speed at which such programs can be put into operation. There is also a limitation as to personnel and funds available to enable strict adherence to an acceptance sampling program. Much still remains to be done in developing the variable or dimensional inspection part of the program. Along with the expansion of the program comes the need for (Slide X) keeping the program "up to date".

With the development of technical "know-how" in glass container manufacture as well as improved mechanical and electronic inspection it is obvious that the quality level will be improved. Increased demands by the user because of higher production speeds and improved methods of product manufacture, necessitate constant pressure on the supplier to produce higher quality ware. Subsequently, comes the need for a constant reviewing of the sampling and inspection program. Re-classification and re-definition of defects, adjustment of quality levels, and improvement of sampling and inspection techniques must be considered at regular intervals.

With this thought in mind, a series of day long meetings was arranged this year beginning in late January and extending through February with each of our glass suppliers. Prior to these meetings an agenda was prepared covering the various phases of the acceptance sampling program. Each supplier had an opportunity to expand on the agenda so as to make it all inclusive.

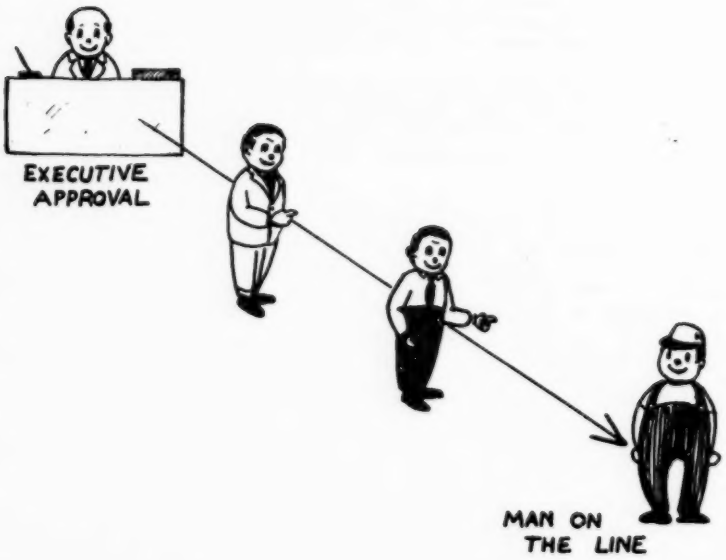
Much was gained from these meetings by preventing some of the problems and pitfalls usually resulting from simple misunderstandings. Each supplier had an opportunity to objectively evaluate the problems of our particular glass container acceptance sampling program.

We still do not have the answer to all our problems--we never expect to. Pitfalls have been avoided and problems have been prevented by a sincere cooperative supplier-user approach to acceptance sampling.

I

WASHBOARDS ?
BIRD SWING ?
CHICKEN ROOST ?
HAIRLINES ?
OVERPRESS ?
ENAMEL
MONKEY SWING ?
CRIZZLE ?
CRACK ?
CUT OFF MARK ?
LAPPED BOTTOM ?
ROCKER BOTTOM ?
HEEL TAP ?
STRINGS ?
TEARS ?
PRESSURE MARK ?
DIMPLED SHOULDER ?

II


EXECUTIVE APPROVAL
MAN ON THE LINE

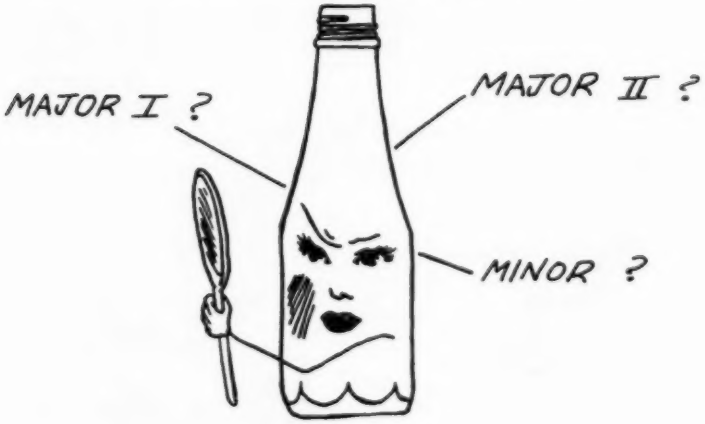
III

AGREEMENT—
BETWEEN

SUPPLIER AND USER

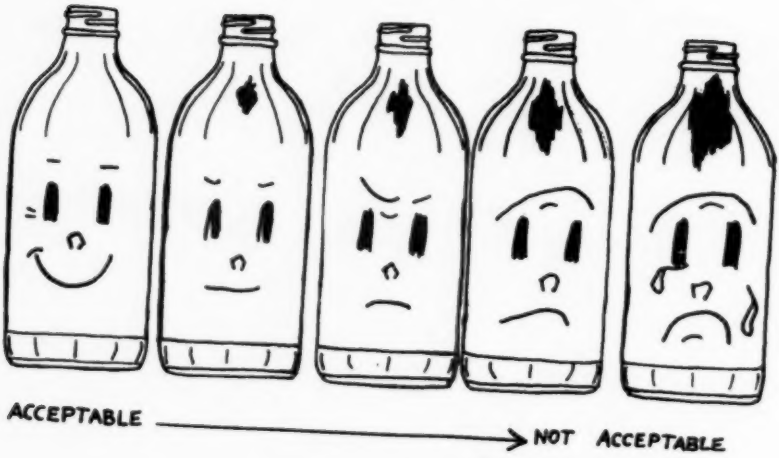
IV

IS THE DEFECT—

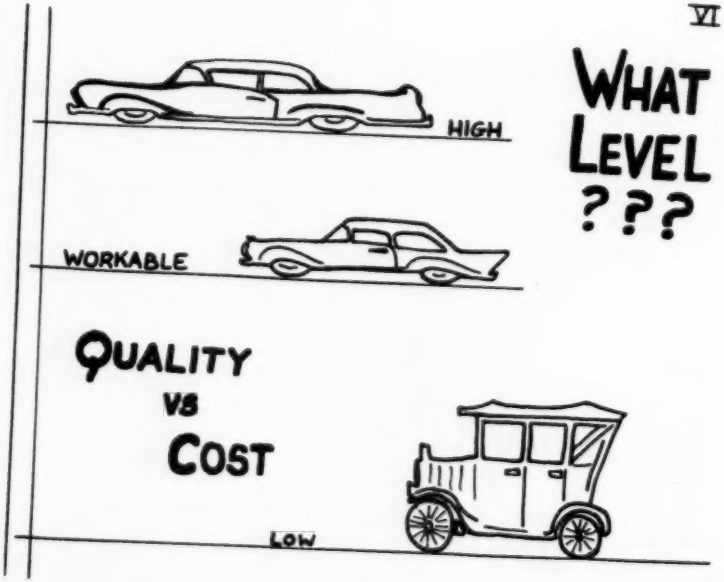


V

LIMIT SAMPLES

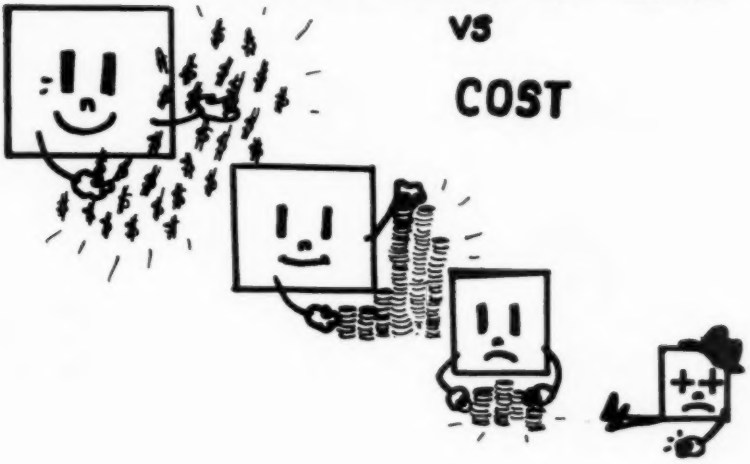


VI



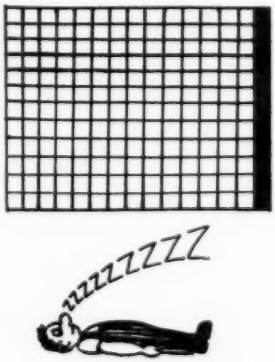
VII

PROPER SAMPLE SIZE VS COST

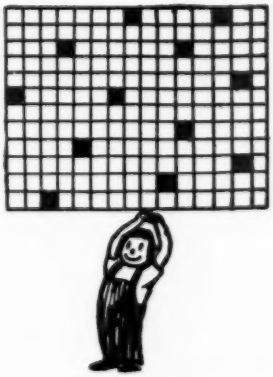


VIII

WHERE AND HOW TO SAMPLE



WRONG – LAZY



CORRECT
RANDOM SAMPLING

IX

SELECTION OF THE RIGHT PEOPLE TO DO THE JOB

X

QUALITY CONTROL APPLIED TO A JOB SHOP

Harmon S. Bayer
Quality Control Consultant

Will it work? Effective quality control programs have been successfully applied to many job shops in various industries. Considerable real savings have resulted from reduced manufacturing losses.

A large screw machine job shop saved over $100,000 within one year. Compared to prequality control records, scrap was down 75%, repair costs were reduced 50%, inspection labor was reduced by 60%, and customer complaints were cut in half. Those complaints still received were of a less serious nature.

A rubber manufacturing concern paid for its quality control program within six months by cutting its high loss scrap items to 20% of the former level.

A tube manufacturer realized an overall savings of $125,000. Quality control studies helped improve machine maintenance and production methods. Customer complaints were down by 75%.

The job shop problem of applying quality control methods and the usual psychological deterrent to attempting a program revolves around short runs and the generally irregular production. These problems are not as large as they appear if job shop management will look at the facts.

All phases of management in job shops have faced this problem and solved it. Sales, engineering, scheduling, tooling, design, maintenance have learned to develop procedures to overcome the short-run handicap. Quality control has also solved the problem of building a system to control short runs in many companies.

The problem is not as complicated as it might seem from a control standpoint due to the repeat nature of job shop business. Historical quality records are kept and applied to repeater items as if no lapse in production had occurred. Usually the problems of the repeat orders are not significantly different from the past quality problems. What has been missing, prior to a quality control program, was an organized method of communicating past experiences in control problems to repeater situations.

The means of organizing a quality control system comes further into focus when it is realized that all job shops seek sales based upon their manufacturing skills and available equipment and processes. Although the production procedures built around a specific order may be different, they are, in fact, only a combination of a few production methods available within the shop. For example, a screw machine job shop may produce one part which requires drill, counterbore and tap, grind, slot, cut-off, wash and degrease. Another part may also be produced which may seemingly have no relationship to the first part because its shape and size may be entirely dissimilar. However, if one analyzes the operations necessary to produce the second part, it may be that the production requirements are very similar to the first. An analysis of all parts which a job shop produces will result in a realization that these parts fall into a small

number of production patterns. This is the key to the organization of a quality control program in a job shop. The control system is built around machines and equipment so that regardless of the current nature of the product, the processes are controlled and the result is that the specific part in production is controlled.

The basic philosophy of quality control must be accepted by job shop management if a program is to succeed. This is the same decision that has to be made by non-job shop companies. Briefly, management must accept the principle that quality is the responsibility of the production departments. The quality control and inspection departments are responsible to develop control programs and procedures, audit the quality, determine the problems, trace their sources, and inform the person responsible so that he may correct the causes of poor quality.

In addition, management must also accept the principle that quality assurance cannot be obtained by attempting to inspect out the bad items at the end of a production system. Quality must be built into the product at each operation and quality control systems must be established to control in-process operations.

Statistical quality control methods are fully applicable and extremely useful to the quality control program in a job shop. Any program contemplated must be built upon the principles of statistical quality control to be sufficiently effective. Some specific problems in the application of the techniques will be discussed later.

The quality control department needs people. The job cannot be done as a part-time responsibility of an inspection supervisor, an engineer, or a lab man. Relatively small shops (50 people) will need at least one full-time quality control engineer. The larger shops need comparatively larger staffs with quality control engineers, technicians and clerks. Of course, shops with less than 50 employees may have to modify the principle of assigning a full-time man, but they will probably find quality control will become his primary, not secondary responsibility. Job shop management may not be so reluctant to hire these people when they realize that a well organized program will not only cover the investment in personnel but pay the company considerable returns.

The functional organization within the quality control department usually consists of the quality engineering section and an inspection section. The quality engineering section develops the statistical engineering services and inspection procedures for the program; the inspection section performs the regular inspection duties, provides inspection labor for routine inspection and control charts and special studies under the direction of quality engineering.

A quality control program for receiving inspection is one of the necessary phases of a complete quality control program. The organization of the program for receiving is no different for a job shop than for any other type of company. There are many excellent references in the quality control literature to help the job shop establish this phase of the program.

In-process control. Job shops have found that they are unable to provide quality assurance by inspecting a product at the end of the

production process. Successful quality control programs have shown that controls must be established at each phase of the operation and systems developed to provide a number of services to production.

Determining the source of quality problems is the major contribution that a quality control engineering section can provide by special studies which analyze machines and equipment. The different aspect that the job shop encounters is that the results must be analyzed not only in terms of the current problem and the specific part number involved, but this knowledge must be so organized to apply the results to future production which may be similar to the problem at hand. In this manner a body of knowledge is built up concerning the quality capabilities of the available machines and processes. By making other departments aware of this information, future designs can be changed, machines and equipment can be improved and maintained and scheduling departments can prevent assignment of specific orders to machinery which the studies have proven to be incapable of maintaining tolerances. These are just some of the benefits to be derived from this activity.

Control charts should be established at difficult operations to aid production in solving quality problems. The use of this technique is no different from straight line production; the clerical procedures must merely be designed around the job shop problem of short runs.

As an example, an average and range ($\bar{X}$ & R) chart was used to help control an overall length of a tube (Illustration A). Past history on a somewhat similar part (although different diameter and length) was used to establish temporary limits on the chart. The first run on January 10 was short and only a few points were plotted. Yet the temporary chart did help get the job corrected (the cut-off saw was adjusted).

When repeat order production started on February 3, the chart was again applied to the job and helped prevent further quality difficulties by indicating a dull saw and a needed adjustment. By the end of the second run, sufficient data was taken to calculate permanent statistical limits for the chart. These proved to be so similar to the temporary limits that they were not changed. Many job shops have found that when a few typical machines are studied for specific operations (such as cut-off) the chart limits connected with various machines and varying sizes are so similar that standard control charts can be prepared and used on similar operations. This is done by reducing all dimensions to the variation from the specification nominal (the midpoint of any specification is defined equal to zero). In many cases the charts are used to reflect the production of several different part numbers, in the order of their production, on a certain machine (or machines) without loss of the chart's ability to detect quality problems. For further information on the application of the $\bar{X}$ & R chart to job shop problems, please refer to the Technical Supplement to this paper.

In-process inspection systems. Control charts and quality engineering studies will help reduce quality problems but an effective quality control program must also develop routine in-process inspection systems to assure control of the product at each machine or process.

Inspection instruction sheets must be developed to give production operators and inspection personnel full knowledge of the customer's

quality demands. Usually, this is where job shops without quality control programs fail in their quality goals. Communications must be effective and complete so that all the information necessary to do the quality job flows from customer contact to production and inspection personnel.

A typical inspection instruction sheet for a job shop (Illustration B) gives details concerning the inspection points at each operation, the specifications, the inspection methods, etc. Posting it at the operation accomplishes the purpose of informing the operator and inspector of the customer requirements.

Preparing these inspection instruction sheets for each operation can be complicated. It is simpler if one remembers a principle stated earlier; most job shops find that although a large number of different parts are produced, actually only a reasonably small number of different production method combinations are used. Based upon this, standard instruction sheets are prepared to cover a family of parts (in Illustration B this is a "Flat Goods" family). These items are typewritten or printed and reproduced on the standard form. The specific items related to the particular part involved are filled in where necessary (in Illustration B these specific items are handwritten). The unused items are crossed out. These clerical methods and many others similar to this must be developed by job shops to simplify the clerical problems created by the large numbers of possible items and short runs.

Jigs, fixtures and gages for checking production must be available at each operation so that the operator and the inspector can check the part at the operation frequently enough to assure good quality.

"First piece" or "first run" inspection must be established to inspect the first few parts immediately after set-up. This should be more complete than a regular inspection, almost layout in character but somewhat simplified. This inspection should quickly detect errors in tooling and set-up. Jobs should not be allowed to run without this type of inspection.

An identification system which identifies all lots of product must be established. Each operator is required to record his identification number, the amount of production and similar items. This makes tracing of the errors possible. It has been found that this identification system is indispensable to an effective program and, in reality, a small investment compared to the dividends received from the system.

Routine roving inspections are made at each operation on the basis of the inspection instruction sheets and standard statistical acceptance sampling plans are used. Please see the Technical Supplement to this paper for further information on these plans. If the lot passes the sample, it is allowed to proceed to the next operation. If it is rejected, it is returned to production supervision for action. This is based upon the previously stated principle that inspection only audits quality; production is responsible for good quality. The production supervisor may elect to return the lot to the operator for sorting or repairs. In any event, the lot cannot be used by the next operation or proceed to the customer without correction by production and resubmission to inspection for approval.

Final inspection or audit inspection should also be accomplished with the use of standard statistical acceptance sampling plans. Job shops will find that this can replace 100% inspection of items and give the management and the customer increased protection at a reduced cost. In addition, as the in-process inspection programs develop efficiently, it obviously becomes unnecessary to audit most items and this effort is curtailed. Lots rejected in this final or audit inspection are returned to the department responsible for the problem for correction.

Complaint investigation. Complaints from customers or from other operations or departments within the organization should be fully exploited to prevent future quality problems. Unfortunately, some organizations merely ask the inspector, "How come that got out?" A quality control program investigates each complaint, determines the source and the responsible production department, requires a statement from that responsible party for the record concerning his actions to prevent future recurrence and projected correction date. In this manner the acceptance of the responsibility of production for quality is gradually developed by the organization and the complaints reach the party that can and should do something about the problem - the man who makes the item.

Quality control will work in a job shop. The techniques are fully applicable; the methods effective. Many job shops have found that the programs pay handsome dividends. The usual deterrent to starting a program is the concern it may not prove effective. It is hoped this paper dispels some of these doubts.

* * *

TECHNICAL SUPPLEMENT

$\bar{X}$ & R charts on short runs where no previous experience is applicable to place limits, use the following effective chart. Assume that the process is just capable of meeting tolerance.

Then

$$\text{Tolerance} = 6\sigma$$

$$\text{Tolerance} = \frac{6\bar{R}}{d_2}$$

Assume sample size of 5 and B/P specification of $0 \pm .005$ (tolerance = .010")

$$.010 = \frac{6\bar{R}}{2.326}$$

$$\bar{R} = .0039$$

Then the $\bar{\bar{X}} = 0$ (specification nominal) and control limits would be:

$$CL_{\bar{X}} = \bar{\bar{X}} \pm A_2\bar{R}$$

$$= 0 \pm .577 \times .0039$$

$$= 0 \pm .0022$$

The range chart limits would be:

$$UCL_R = D_4\bar{R}$$

$$= 2.11 \times .0039$$

$$= .0082$$

$$LCL_R = D_3\bar{R}$$

$$= 0$$

The significance of the chart is as follows: Place the chart on the job immediately. The range chart out of control indicates the machine or process is not capable of maintaining the operation within specifications. Out of control on the $\bar{X}$ chart indicates need to adjust to maintain specifications. If process is better than specifications, the $\bar{X}$ chart may cause some overcompensation since the limits will be narrower than is necessary for control within specifications.

On $\bar{X}$ & R charts for a process where the capability is better than the specification, it is advised that modified limits calculated as follows be used:

$$UCL_{\bar{X}} = \text{Upper Spec} - M\bar{R}$$

$$LCL_{\bar{X}} = \text{Lower Spec} + M\bar{R}$$

$$\text{where } M = \left(\frac{3}{d_2} - A_2\right)$$

Statistical sampling plans. Because of the confusion of parts and production, it is necessary in a job shop to make a basic decision in the use of statistical sampling plans to limit the use to a few standard plans. For example, the decision might be made to standardize the in-process plans as follows:

Defect Classification	AQL	Sample Size	Acceptance Number	Rejection Number
Critical	0.25%	50	0	1
Major	1.0%	15	0	1
Minor	6.5%	5	0	1
Incidental	10.0%	3	0	1

Similarly standard plans would be developed for audit and receiving inspection. Through this method of standardization the training of inspectors and the carrying out of procedures can be simplified. Inspectors can be taught a simplified routine connected with a few sampling plans, and can concentrate on the confusion created by the constantly changing inspection instruction sheets. The quality control engineer and the specification engineer, on the other hand, can place a particular defect in any of the standard AQL's or in some cases, if they do not strictly apply, change the tolerance to make it applicable

to any of the standard AQL's. For example, a .010" tolerance with a 3% AQL is approximately comparable to a .012" tolerance with a 1% AQL.

* * *

REFERENCES

1. ASTM Manual on Quality Control of Materials, STP15-C, American Society for Testing Materials, Philadelphia 3, Pa., January 3, 1951.

2. Bayer, Harmon S., How to Use $\bar{X}$ and R Charts Effectively, American Society for Quality Control, National Convention Transactions, 1955.

3. Bayer, Harmon S., How to Determine Quality Capability of Machines and Tools, Tooling and Production, May-October, 1954, Huebner Publications, Inc., Cleveland, Ohio.

4. Bayer, Harmon S. and Engelgau, Ray F., Selling Management a Quality Control Program, Tooling and Production, April-May, 1956.

5. Bayer, Harmon S. and Engelgau, Ray F., Quality Control for Automatics, Steel, February 18, 1957, Penton Publications, Cleveland, Ohio.

6. Duncan, A. J., Quality Control and Industrial Statistics, Richard D. Irwin, Inc., Homewood, Illinois, 1953.

7. Grant, E. L., Statistical Quality Control, McGraw-Hill Book Co., Inc., New York, 1952.

8. Statistical Quality Control and Acceptance Sampling, Ordnance Inspection Handbook, ORD-M608-9, U. S. Government Printing Office, Washington, D. C.

9. MIL-STD-105A, Sampling Procedures and Tables for Inspection by Attributes, September 11, 1950, U. S. Government Printing Office, Washington, D. C.

10. Statistical Research Group, Columbia University, Sampling Inspection, McGraw-Hill Book Co., Inc., New York, 1948.

ILLUSTRATION A

AVERAGE ($\bar{X}$) & RANGE (R) CHART - FOR OVERALL LENGTH

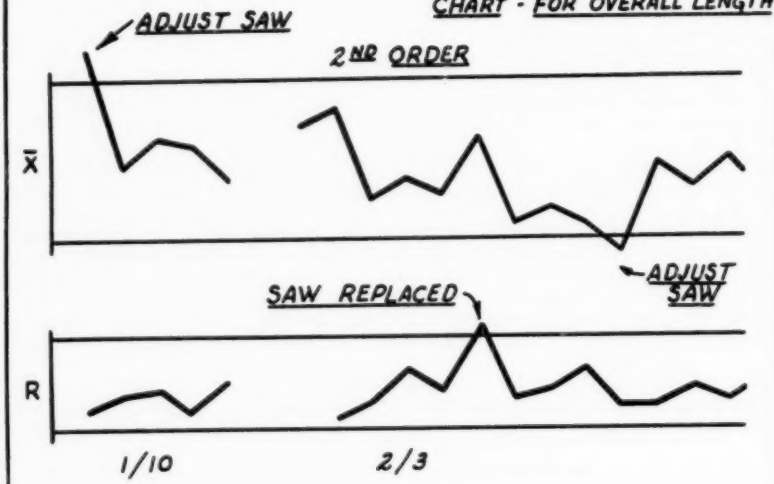


ILLUSTRATION B

TY CONTROL DEPT.

MATERIAL INSPECTION INSTRUCTION SHEET
FLAT GOODS

PART NO. 12

ON DEPT. PRESS

R CUSTOMER XYZ COMPANY PART NAME SEAL

WING DATE 1/11/55 LATEST CHANGE LETTER A

CRITICAL 0.4 MAJOR 1% MINOR 10% INCIDENTAL 15

CHARACTERISTIC	SPECIFICATION	INSPECTION METHOD
Major		
Gas Burns	None	Visual
Non-Fills	None	Visual
Gauge	.120" +.005 -.010"	Micrometer
Length	28.50 ± .060	Rule
Width	1.50 ± .020	Rule
~~O.D.~~		~~Micrometer~~
~~I.D.~~		~~Indicating Gage~~
Minor		
Dirty Mold	None	Visual
Cure	76 - 80	Durometer

THE DESIGN & ENGINEERING OF TEST EQUIPMENT FOR COMPLEX ELECTRONIC PRODUCTS

Michael Gebrian
Sperry Gyroscope Company

Introduction

(1) Perhaps it may be more correct to paraphrase the title, "The Design & Engineering of Complex Test Equipment for Modern Electronic Products". Electronic and electromechanical products are all complex, to varying degrees, and always have been; however, the recent rapid advances manifest in the design of prime military products in order to meet exacting accuracies present a challenge to the test equipment designer and a problem in the maintenance of quality control.

(2) It is not the intent of this paper to present a detailed technical exposition of test methods and procedures commonly employed in the industry; numerous books and pamphlets on this subject have been published by military and civilian agencies. Rather, an attempt will be made to highlight some of the difficulties faced jointly by engineering, manufacturing and quality control in the production and acceptance of primary military equipment, and to suggest possible means for circumventing these difficulties.

(3) Briefly, these obstacles fall into several general areas:

(a) The effect of the measuring tool on the quantity to be measured.

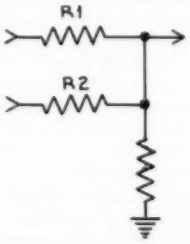
(b) The need for close coordination of the two major design concepts - product and test equipment.

(c) Administrative control of test equipment design.

(d) The goal of special test equipment design.

The Effect of Measurement

(1) The inspector who measures the O.D. of a shaft with a micrometer is probably unaware that the reading he extracts is in error by the amount of distortion the shaft is subjected to because of the pressure of the caliper. If he is aware of this, he can rightly discard this error; it is either unmeasurable or too minute to be considered. However, the test methods engineer faced with the problem of measuring torques of $\frac{1}{4}$ dyne centimeter, voltages and frequencies in the order of fractions of one percent, and ratios of a resistive summing network to .01% cannot disregard the effect of the device he uses or the method employed upon the quantity he observes. The error contribution may well equal or exceed $\frac{1}{2}$ the tolerance spread of the particular parameter; thus, he cannot justifiably accept or reject the unit under test. As an example, consider two resistors, R1 and R2, which form a summing network as shown:



It may be required that each resistor has a tolerance spread of 2%, which can be measured, but that they be matched to achieve a ratio of 1 ± .01%; ie, the ratio R1/R2 = 1 ± .01%. Selection of matched pairs by ohmic resistance measurement is not practical; in the worst case, the individual values of R1 and R2 would have to be determined to an accuracy of .005%; such accuracy is not obtainable with present ohmeters, including the digital type. Another approach is necessary: one that is feasible is a nulling method employing the resistors and a precision ratio transformer in a bridge circuit. Even then, the error contribution of the transformer may have to be taken into account.

(2) It is not enough, however, to determine a technique which will minimize the effect of the measuring tool; it becomes necessary to evaluate the extent to which product accuracy is degraded by the type of wiring employed between the test fixture and the unit under test. In many cases, observed readings are appreciably affected by wire type, electrostatic and/or electromagnetic shielding, lead dress, and by the grounding method employed.

(3) The simple example stated above is illustrative of the extreme care which must be exercised by the test fixture designer in his approach toward methodizing a particular test operation on either a system or a component of a system; on the other hand, it is incumbent upon the product designer and/or the engineer responsible for the preparation of the test specification to not only keep up to date with "state of the art" test concepts and instrumentation, but to include errors introduced by monitoring equipment along with product component deviations in his calculations of nominal values and tolerances.

Close Coordination of Product & Test Fixture Design Concepts

(1) It is a truism that the ideal is not attainable. The present highly competitive state of the electronics industry in the area of government contracts dictates that beside cost, quality is an important factor in obtaining sales. The key to quality is performance in terms of reliability and accuracy, attainable to the extent that the basic design approach of the prime product and the selection of its components are sound. The engineer must apply particular care to avoid the dilemma of uncertainty which may occur when his prime product fails to meet required tests. He may not be able to determine conclusively whether the design of his product is faulty in some respect, the test specification is invalid or unrealistic, or whether the test philosophy employed has degraded the performance of the system.

(2) The performance of an electronic product can be theoretically calculated by determining the root mean square value of all the errors produced by the components which make up the system, provided these errors are independent of one another. In the process of establishing the test specification nominal values and tolerances of the system and its sub-assemblies, the product engineer must include the effect of the test device, and must provide the necessary access points for testing. This is particularly applicable where wire length may contribute to a deviation from the nominal value as well as the monitoring device itself. By administrative fiat, the effect of the test fixture upon the product being tested may be discounted if it contributes no more than 20% to the allowable tolerance spread of the specific test function. The application of this rule is a good yardstick for the designer of test equipment, but the product engineer must take into account the limitations it may impose. For example, if a voltage is to be metered to an accuracy of .05%, the instrumental error is limited to .01%; this precision is not practicably attainable for the measurement

of absolute values. An indirect approach would be necessary, perhaps by a determination of a relationship of that voltage with respect to another parameter.

(3) As a general rule, it is good practice to simulate the actual environment a product subassembly, undergoing test, might encounter if it were installed in the next higher assembly. That is, the load that a product amplifier will "see" in the next assembly may well be incorporated as part of the test equipment. Frequently, the product engineer will specify that the test of a component be conducted with its associated product mates, or equivalent. However, this practice must be analyzed in every instance by the test methods engineer to prevent reductio ad absurdum. For example, making a closed loop test of a dissociated servo amplifier by installing all the other servo components in the test fixture would be pointless; it may serve to evaluate the basic product design concept, but from a production viewpoint, it only serves to "test the tester". In addition, the concept of system simulation may lead to an element of uncertainty in the use of a product part as a test fixture load; application of the 20% rule may serve well to reduce the effect of the measuring tool, but may not reflect true system performance.

(4) The high order of precision required in modern electronic gear demands extremely close collaboration between the product engineer and the test equipment designer; each must be familiar with the design problems preculiar to his colleague's area of responsibility. Test specifications must be valid and realistic, while the associated test equipment must conform to the specification requirements. The two functions cannot be divorced; the two principals constitute a technical team with the common goal of producing a verified complex product.

Administrative Aspects of Test Equipment Design

(1) The art of special and general instrumentation is hard pressed in keeping up with the fast pace of technological progress of military and civilian electronic products, especially in the field of inertial navigation. As a result, it has become a highly specialized area of test philosophy and techniques in which the test equipment engineer, test engineer, or test methods engineer, whatever his title, represents the technical mainstay. His is an important role in the control of quality and economy, for a successful test fixture design presupposes a thorough familiarity with the product, with engineering practices, manufacturing methods, capabilities and limitations, and with the progress of electronic instrumentation.

(2) The question as to whether administrative control over test equipment design should be an engineering function or a quality control responsibility is a debatable one. On the one hand, it can be argued that coordination of the product and test design approaches can be better effected under central control, and will result in economy by reducing duplicity of effort. In this sense, duplication is avoided by repackaging the breadboard test set-ups initially employed in product design evaluation. In addition, compatability between design evaluation and production test results is more easily achieved since the method of test is the same in both cases, with only the added factor of human engineering applied to the factory test fixture. Therefore, engineering control is warranted, for the importance of close liaison between the two design functions has been emphasized before.

(3) On the other hand, there are several reasons to support the contention that test fixture development is a quality control function. Of course, a basic assumption must be made that the

necessary technical skill is available in either case, so the question is purely one of administration. For one, test results of a product or its components are as much a measure of quality as are control charts for a machining operation or methods for statistical sampling. Control over these results can be exercised by monitoring all phases of the manufacturing process to the final stage of customer acceptance. This will depend for the most part on the performance rating of the system, determined by testing to a specification with special test equipment. On a production basis, this is a manufacturing function; although quality control is geared to manufacturing processes, it should play a positive role in the preparation of the final test specification. This is important, since this document is the basis for design of special test tooling. It follows that quality control should carry on with the task of producing the test equipment.

(4) Here, incidentally, is one phase of quality control activity which does not require much selling to management. In its final state, an electronic product lies somewhere between engineering's ideal version and manufacturing's practical approach to the ideal. In terms of performance, its quality will vary between these limits, yet may stay within the acceptable region. In terms of cost, however, the difference might be prohibitive the closer the ideal is approached. It is in this activity that quality control can insure a proper balance between cost and performance by having a voice in the preparation of the test specification, and by designing test tooling with this goal in mind. It is debatable whether this objective would be attained as effectively if a single agency were responsible for product design and engineering, the generation of test specifications for that product, and for the production of factory special test equipment. Generally speaking, engineers, collectively, tend toward idealization; manufacturing leans toward the practical. Quality control is an effective buffer between the two divisions in the field of testing.

The Goal of Test Equipment Design

(1) The characteristics of properly designed test gear are straightforward and unimposing, yet, like the proverbial stitch in time, usually achieved by hindsight. This is true especially if funding and time are restrictive; yet they are essential for minimizing potential product difficulties. Briefly, the goal is fourfold:

(a) Simplicity is important. It is not to be inferred that the design is simple in concept; on the contrary, modern electronic test equipment can be quite complex in detail. By simplicity is meant uncomplicated in operational use, and designed with a view to reduce instrumentation and/or circuitry to a minimum; in effect, to discard the Rube Goldberg approach in favor of the more direct. Cumulative errors increase in proportion to the amount of devices appended to the unit under test.

(b) Reliability is becoming more and more necessary, in keeping with stringent reliability requirements of guided missiles. Inherent in reliability is the element of accuracy, of course, but an equally important feature is durability. The designer is often tempted to strive for an aesthetic result in a particular "black box" which may perform a required function in an elegant manner. This may impose a problem of maintenance and calibration, however; test fixtures are in constant use in production testing. For example, a device which must produce a series of discrete timing marks at specified intervals could be done very elegantly by a bank of electronic flip-flop circuits, but a mechanical gear and cam setup will be more stable and long lived.

(c) Accuracy is self-evident, as a characteristic of dependable test fixtures. Many dilemma producing questions about the performance of a system can be avoided by the proper choice of circuitry and instrumentation used in the test equipment. Included in this is the element of repeatability of observed readings.

(d) Flexibility is the keynote of design. It is becoming more and more apparent that the number of emergency and normal changes to an electronic system vary directly as its complexity. One thousand per month is not unusual, and many of them affect the associated test equipment. These changes are a bane to the designer's peace of mind, and frequently the subsequent fixture modifications are extensive enough to cause a delay in production. Flexibility is a most desirable feature, and may be realized by careful thought to switching, choice of instrumentation, and by the judicious use of programming devices such as patchboards or punched card systems. Adaptability also contributes to ease of maintenance and calibration necessary periodic operations for the maintenance of quality.

Conclusion

(1) It is beyond the scope of this paper to discuss automation concepts and techniques as applied to test equipment. Designing to this end represents the ultimate in complexity, and is difficult to accomplish unless product drawings are relatively frozen. Semi-automatic or manually operated electronic test equipment is intricate enough, especially in the field of inertial guidance or missile systems, which are undergoing rapid technological evolution. Support equipment is big business, however, and the trend toward automation for field use is a question of time and demand, limited only by the test engineer's ability to keep pace with new developments.

(2) The principles outlined in the preceding paragraphs are fundamental, yet they present a challenge to engineering and quality control. To draw ananalogy: the contriver of the weapon of defense faces the prospect of not being able to cope with a new offensive weapon. The test engineer's nightmare is to be confronted with a new product which cannot be tested conclusively.

QUALITY CONTROL'S OBLIGATION TO MANAGEMENT AND CUSTOMER ON RELIABILITY OF COMPLEX WEAPONS

J. W. Young
North American Aviation, Inc.

The subject of quality control's obligation to management and the customer on reliability of complex weapons is very broad and can be discussed from many viewpoints. I will consider our obligation to management and our customer as the same, since it would seem that management of a company would have the same feeling toward delivering reliable products as the customer would have in receiving them. Also, I will consider management as top level management in the company and our customers as the military services, as evidenced by our subject and two of our panel members.

There are many ways that companies have changed or supplemented their organizations to deliver more reliable products to the services. I do not feel there is an ideal organization for reliability or any similar aspect of a complex weapon which can be introduced in a given company. Organizations are often decided by the capabilities of available personnel, as well as the overall company organizational structure.

Regardless of where the group or groups identified with reliability are positioned in the organization, there must be the desire and enthusiasm of top management and all departments involved to accept the concept of reliability and work toward these reliability goals. It is apparent that engineering, quality control, manufacturing, and purchasing are the departments who can contribute the most in delivering reliable complex weapon systems.

While I feel that engineering, by the very nature of its work, will always carry the greatest share of the load and make the greatest contribution to reliability, quality control is also a major contributor.

It has been stated that quality control is interested in all activities that are directed toward the production of usable and reliable products at a minimum overall cost. It is obvious that a quality control function in an organization cannot have such broad responsibility. However, quality control's usually recognized responsibility naturally requires it to be interested in most activities in the company which influence the delivery of reliable products.

The very nature of quality control and inspection lends itself to the recognition and need to improve features in the products which experience, good judgment, or statistics indicates a change should be made immediately or at an early change point. In addition to assuring that a quality product is being delivered, we must get data back to engineering, manufacturing, and purchasing in a factual form which will help them, and in cases where action is needed, follow through to be certain that it is taken.

I believe there will be very few instances where another department will not welcome and encourage information and recommendations from quality control if they are properly presented and which will help that department to do a better job. To make this contribution, quality control must command the respect of engineering, manufacturing, and pur-

chasing. We will only get this respect by developing our people to appreciate their responsibility in this regard and also to understand the problems and nature of the other departments' work.

Too many times, I have heard engineering personnel say that they had to make investigations or gather data themselves rather than rely on quality control because they did not think quality control had personnel who could do the job. This may or may not have been the case, but it was apparent that quality control had not been aggressive and done the job they could have done or possibly in the way they should.

If we are going to build complex weapons with the reliability that our customers are asking, quality control must be an aggressive, alert, and technically minded organization. This is an approach the head of the quality control organization must promote, not by words, but by actually showing other departments the contribution they can expect.

As our weapon systems become more complex, the prime contractor and second tier contractor will be getting many more systems and functioning parts from suppliers. These larger and smaller suppliers in many cases do not fully understand the conditions and environments under which their products will be expected to operate. Also, in many instances they do not appreciate the quality level and reliability that the complete complex weapon requires, or really appreciate the job they have contracted to do.

I feel that relations with the supplier is a very important phase of reliability where quality control has a principal role, and where there is real work to be recognized and accomplished. We must do a great deal more to evaluate the past performance of suppliers and survey their capabilities to build reliable products before they are given a contract. I do not mean just compiling suppliers' rejection records or rating suppliers on past performance.

After a supplier has been recognized as being capable of doing the job and has received a contract, we must work with him and recognize trouble areas and bring engineering and manufacturing in to help the supplier as early in the contract as possible. We must bring the supplier into our plant for quality and reliability symposiums and have quality control, engineering, purchasing, and manufacturing participate, and encourage the supplier to feel a part of the team.

There are many items which are purchased to a design objective specification written by engineering. Oftentimes, the supplier designs the part and makes decisions which the contractor's past practice and experience have shown to be poor. When engineering reviews a supplier's design before parts are made, quality control can contribute materially with experience and test results they have gained from similar parts in the past. These, again, are areas where quality control must make their contribution to the economic production of reliable products.

Quality control is in the position to observe deficiencies originating in all phases of the production of complex weapons. In this continually changing business of producing complex weapons, management expects quality control to develop an organization which is progressive, forceful, and effective in contributing to the reliability of complex weapons.

THIS IS INSPECTION AUTOMATION

W. W. Spencer
General Electric Company

Is automation, inspection automation, in your future? If one is to judge by the tremendous surge of interest and evident anticipation with which most people react to this word "automation," it certainly is. So that such interest and anticipation may be satisfied, let us examine this matter of inspection automation for the purpose of putting it to work in our business.

Inspection automation is that portion of automation that relates to the taking of measurements, processing the resulting data and feeding back the results for control. We can hereafter refer to process control automation whenever we wish to indicate inspection automation. This leaves us with process automation as the other side of the coin. Together process automation and process control automation make up total automation.

Process control automation is growing rapidly in importance to all of us. Many experts agree that in the next ten years larger and larger amounts will be spent by industry for automatic control of manufacturing. In fact by 1967, today's amount spent will double and may very well triple. Thirty percent of all manufacturing equipment dollars will be spent for quality control equipment and fifty percent for all control equipment. Automation certainly is in your future.

Some refer to the age of automation as the second industrial revolution. Automation is not a revolution, but rather it is an evolution. Neither is it new. In 1784 Oliver Evans built a completely automatic flour mill near Philadelphia. This mill was truly automatic since there were no workers in the mill. True automation may be defined as continuous automatic production. There are, therefore, several steps from manual operation to full continuous automatic production. These steps start in the manual area and move through the mechanization area to automation. There will be many justified instances in our business where the nearest to automation some processes will come will be mechanization.

For instance in 1661, in Danzig, Germany a mechanized loom was built that wove as many as four to six webs with variable complex patterns. This loom and its successors have produced much wealth. Whether it should ever be systematized into a continuous automatic line from fleece to cloth bolt is a question only a sound economic analysis could answer.

Other processes will never be automated because we are unable to identify and define the specific quality criteria we should be measuring to produce automation. In the petro-chemical industry we have a high degree of mechanization based on measurements like reaction temperature and pressure and material flow, but due to our inability to accurately identify the criteria for product mix, such as fuel oil, kerosene and octane gasoline and the quality criteria of the products, we cannot measure them and feed back data for complete automation, which includes automatic process control.

In mechanical processing or manufacturing we do not have this limitation.

Instead, we do have a unique opportunity to reach out towards process control automation. Since we can clearly define and measure most quality criteria in a machining operation we can have a process control mechanization as a definite opportunity.

The temptation in machining is to become trapped by thinking only of individual operations and not the overall system. We must systematize any proposed mechanization and justify it only in the light of what is best for the entire process of raw material to finished product.

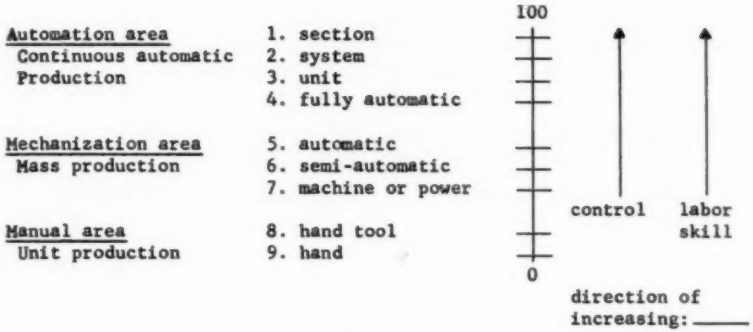
Mechanization and automation may be, in some cases, sure-fire ways of achieving sudden death, of putting ourselves out of business. A profit, an economic benefit, must accrue to the total process in order to justify whatever degree of automation we attempt. Sometimes this benefit comes from cost reduction, increased production, better quality or just being able to do something we could not do in any other way. Thus a properly evaluated and systematized mechanized or automated process, including process control, will be profitable.

In cases studied, it has been found that process mechanization is very profitable but at times it has been handicapped by little or no process control automation. As a result, process control has become a bottleneck. The truth of this is borne out by your own knowledge of mechanized processes where material or product is removed from mechanized production for hand inspection and returned to the continuous operation. On one cylinder block line where a mechanization profile has been drawn, there are twenty-eight process operations and eight process control operations. Twenty-four process operations are mechanized above the hand stage but only one process control or inspection operation is in this category. This is a typical situation.

It would be helpful to have some way of measuring our manufacturing so that we might clearly and simply see the true situation. Just how much of a bottleneck is process control because of a lack of mechanization? Since automation is an evolution, we should be able to set down the principal steps one climbs from bottom to top and match our attainment; i.e., how far up the ladder we have to climb for process mechanization and process control mechanization - Fig. 1 shows one way we can picture the evolutionary ladder for process automation.

Fig. 1

PROCESS AUTOMATION
SCALE OF PROGRESS



Note as we progress up the ladder we require an increase in the amount of control (process control) and also labor skill. This means that unless our process control mechanization keeps pace it becomes a more and more severe bottleneck.

Now to construct a corresponding ladder for process control automation. Our process control language differs somewhat from our process, so Fig. 2 should help us better to visualize what we will mean by the terms of our process control automation scale of progress.

Fig. 2

ANALYSIS OF PROCESS CONTROL MECHANIZATION LEVELS

Level	Power Used	Measurement Equip. Used	Information Obtained
Manual			
gage	hand	fixed and adjustable gages	go-not go
gage assisted	hand	indicating and recording gages	variables
Mechanization			
semi-mechanized	mechanical assist	continuous reading gages	trends
mechanized with classification	mechanical	sorting machines	simple post-process decisions - on and off control
Automation area			
automatic with step control	electrical	continuous measuring machines and devices	simple post-process decisions - on and off control with step adjustments
automatic with continuous control	electrical	continuous gaging	multiple in-process and post process decisions and adjustment
automatic with full system control	electronic	multiple continuous gaging	Pre-process, in-process and post-process decision with prediction and adjustment
automation section control	electronic	multiple continuous gaging	complete interlocking systems with full decision and total control

The manual area starts with the use of simple hand gages which are either fixed or adjustable and give go and not-go information. The first upgrading would be to indicating gages which give us information in terms of variables.

Following this is mechanization. Then it becomes possible to upgrade so that information can be obtained to indicate process trends. Following this post-process decisions can be made that will exercise on-and-off controls.

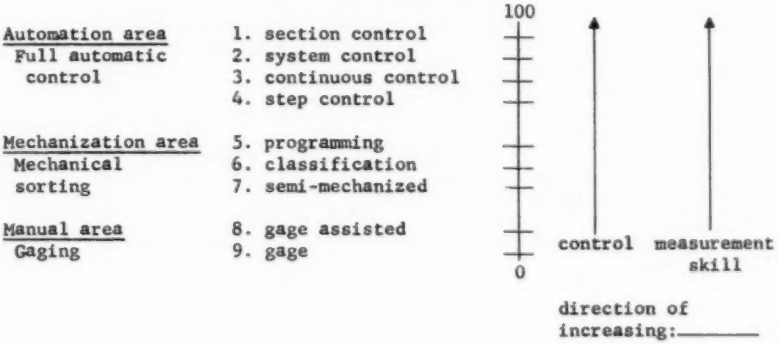
And last, the automation area, where there is a gradual evolution from on-and-off controls with step adjustments through in-process guidance and adjustment of the actual process itself to pre-process measurements with prediction control.

The final step is the tying-together of individual machines with interlocking controls into blocks, lines and ultimately a factory.

From Fig. 2 we construct Fig. 3

Fig. 3

PROCESS CONTROL AUTOMATION
SCALE OF PROGRESS



Again, as in process automation, the start is in the manual area and progresses upward to full automation with increasing control or skill also in the same upward direction.

By combining on a scaled grid the scale of progress for process automation (Fig. 1) with the scale for process control automation (Fig. 3), we can now actually indicate the status of automation for (a) process and (b) process control. We have plotted lines A and B from typical operation information.

Fig. 4

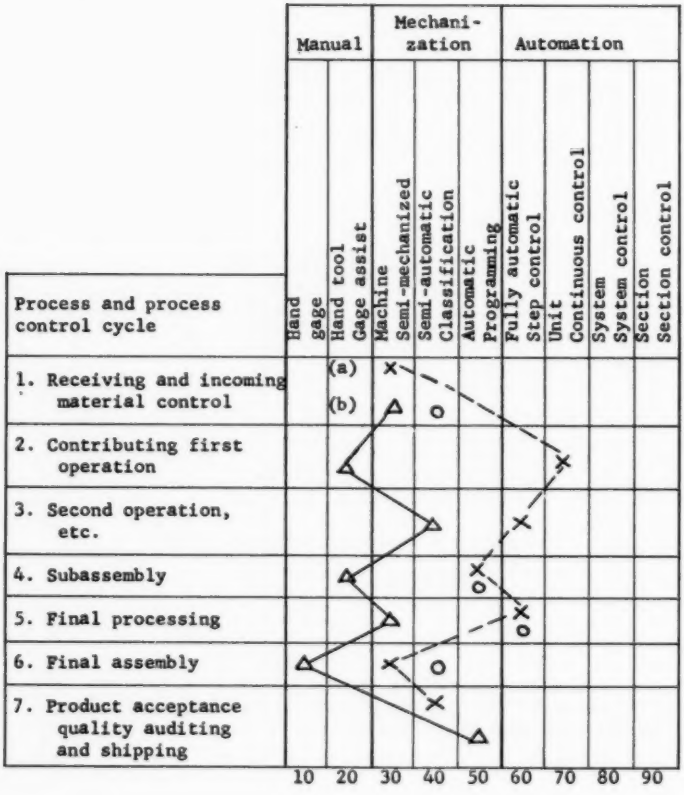


Fig. 4 permits us to recognize the relationship between process and process control automation for each type of operation or point in the manufacturing cycle. It also permits us to average the difference between the two through numerical indexes. In the case shown in Fig. 4 they would be:

process mechanization = 48.6 average index number
process control mechanization = 28.6 " " "

We clearly see by either relationship that there is a long way to go in automation - especially in process control automation. We can also use this diagram as a basis for determining our greatest need and area of greatest potential improvement.

However, before proceeding to correct these unsatisfactory relationships let us look at some other helpful tools for more completely analyzing how to take corrective action.

First, answers to these questions should be found:

(1) What are the quality characteristics that we can measure and use the information for controlling the operation?

(2) Are these measurements properly organized in respect to the total manufacturing system?

(3) Can we make available "sensing" elements for these measurements?

(4) Will the sensing elements feed out useable signals for the control system?

(5) Is this mechanization economical?

With this information in preliminary form at hand, we construct four definitions applying to the specific operation we wish to upgrade through mechanization and automation. These are definitions of:

(1) The measured quality criteria
(2) The transducer or sensing device
(3) The data processing
(4) The data feed back

The next step is to break down the automation of our operation into progressive operational steps or elements and mesh these with our definitions and decide the degree of automation desirable for each element. In many process control areas these elements might be:

(1) Transport
(2) Program
(3) Calibrate
(4) Position
(5) Measure
(6) Record
(7) Analyze
(8) Decision
(9) Feed back
(10) Unload
(11) Sort

Finally we go back and review our first five questions and four definitions and integrate our total knowledge into one single answer - the degree of automation we should have for this operation. If we wish we can make this three step analysis - questions, definitions and operational elements - and from it determine a point for process control automation on our diagram. If we plot this point as a circle on our scaled grid (see fig. 4) we find our total story of process control automation available to us in one pictogram, where we note the potential process control average index number becomes 50.0. This index is just slightly higher than the process automation itself and should represent no bottlenecking effect. Fig. 4 now is a complete diagram against which we can plan our future. At this point we should note any upgrading of the process automation beyond that shown which should cause us to re-evaluate our diagram.

The average example we have picked here may not be too far from any specific operation you may desire to analyze. Experts tell that only 30% of our industrial processes will be automated in the next ten years. Many of these processes will only be mechanized as there will be no economic reason to do more, or perhaps technical know-how will be lacking. In any event much can be done with just what we have available today. The accompanying photographs show us what is being done. As you will note, these photos show equipments that fit the classification shown in Fig. 2. Such equipments are available from most gage manufacturers.

The examples we have just seen might be called examples of dynamic gaging that help produce process control automation. How many dollars have you budgeted for dynamic gaging? Have you analyzed your need for process control automation? Are you knowledgeable of the dollars savings available to you through process control automation? When you answer these questions for yourself, you then can know and say, "For me this is the inspection automation need." Investigate and find out what your inspection automation need is.

A MALTING COMPANY LOOKS AT SPECIFICATIONS THROUGH STATISTICS

Robert H. Bauknecht and William G. Artis
Rahr Malting Co.

INTRODUCTION

Our company is celebrating its 110th anniversary this year and the problem of Quality Control probably began the first time a consumer asked us to "make the next wagon of malt the same as the last one". No doubt an effort was made to comply with this request wherein our old time artisans soon recognized that the solution hinged on not achieving exactness but on reducing inexactness to an unrecognizable minimum in terms of the consumers evaluation.

The transition from Quality Control to Statistical Quality Control is in most cases a gradual movement. The first known application in our industry dates back to the early 1900's when William S. Gosset, Head Brewer of the Guiness Brewery in Ireland, writing under the pen name of Student, reported some interesting deductions based on a statistical treatment of his data. Some 50 years later, statisticians in the malting industry and its associated field "brewing" are again appraising the problems of their endeavors through the revitalized vision of Gosset with more and more enthusiasm for the confidence it gives them in expressing their ideas in an orderly and effective manner. What happened to the theory of statistical treatment of malt house data or why it was not more universally applied in its creators field during all this time is somewhat of an enigma. We have S.Q.C. working in almost every phase of our operations, in some cases in the final status based on current needs and in others, just starting. This paper is pointed specifically at some of the problems we have incurred in meeting specifications and showing how statistics were used in resolving the difficulties. We hope that in coming years we will have the opportunity to present the results of some of the other applications.

DESIGNING SPECIFICATIONS

The existence of a Quality Control Program implies that some type of evaluation has been used to arrive at a set of suitable standards or specifications by which quality is realistically measured. It also implies that conformance to the specifications can be appraised in some manner.

A specification, quite simply, is a definition. For the manufacturer, it defines a quality level which his incoming raw materials must meet so that the subsequent processing steps can be maintained in a predictable and controlled state. More precisely, the specification states knowledge of a permissible range or variation, and that when this range is violated it is possible to measure an undesirable effect either in units of cost or product quality. For example, 140-12 = 128 is the only solution to this problem in mathe-

matics and 132, 125 or 122 are immediately understood to be wrong. No such simple accuracy is available in a malt analytical factor such as diastatic power. Who could say, with four different laboratories finding 122, 125, 128 and 132 degrees Lintner on the same sample of malt, which value is correct? And of even more importance, would it be possible to correlate these four analytical values with any difference in the final product? Yet specifications such as, diastatic power 126-130°L preferably 128°L are often encountered. In other words, there is no cognizance taken of the variation involved.

To get a simple idea of what this variation looks like, take a quantity of malt and divide it into two portions, A and B. Further, divide A and B into about 20 smaller samples. Pick out the malt factors which are considered important and obtain duplicate analyses on the 20 A samples (the first result being called the original and the second the duplicate) and single analyses on the 20 B samples. Then the data are tabulated in histogram form to obtain a frequency distribution diagram.

The results of such an experiment for diastatic power are shown in Figure 1. The diagram at the top represents the total variation or range one could expect from continuously sampling and analyzing a quantity of malt. To obtain the lower diagram for analytical variation, take the results from the A samples and for each sample subtract the duplicate result from the original result and carry the plus or minus sign. The distribution diagrams present a fair picture of the variation one can expect. However, a more precise mathematical definition of the variation can be determined by calculating the standard deviation and distribution curves. The equations are

$$\text{S. D.} = \sigma = \sqrt{\frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N - 1}}$$

and $$Y = 0.4 \frac{Ni}{\sigma} e^{-\frac{1}{2}\left(\frac{(X)}{\sigma}\right)^2}$$

where N = Number of values

i = Class or cell interval - width of abscissa value for each group - here 1°L

Figure 2 shows the frequency diagrams from Figure 1 with the calculated frequency curves drawn in. By definition, 68% of all the data will fall within $\pm$ 1 S.D. of the average, 95% within $\pm$ 2 S. D. and 99.7% within $\pm$ 3 S. D.

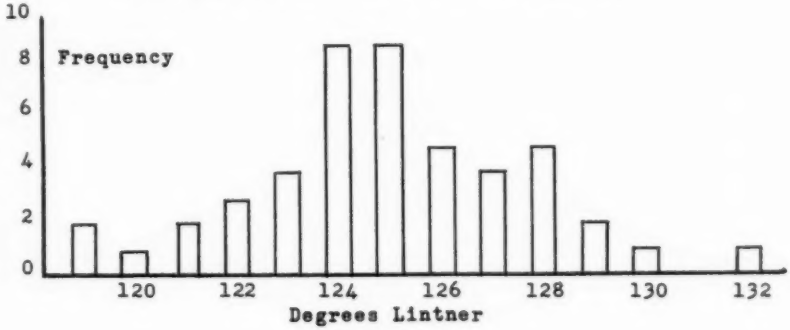
By actually conducting such an experiment, the rudimentary concepts of variations in malt analytical factors are often developed for the first time. Following this, reference to many excellent texts provides information on the de-

FIGURE 1

PRODUCT AND ANALYTICAL VARIATION

MALT FACTOR: DIASTATIC POWER

Frequency Distribution For Product Variation
One Malt Sample Divided For 48 Analyses



Frequency Distribution For Analytical Variation
Difference Between Duplicate Determinations
One Malt Sample Divided For 24 Analyses

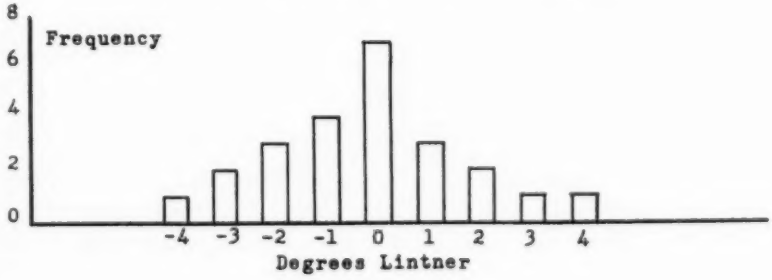
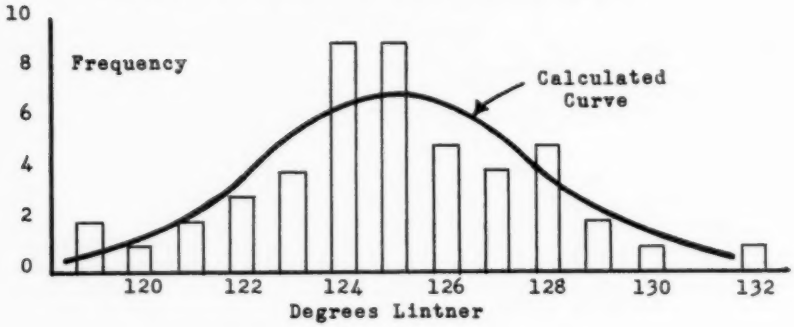


FIGURE 2

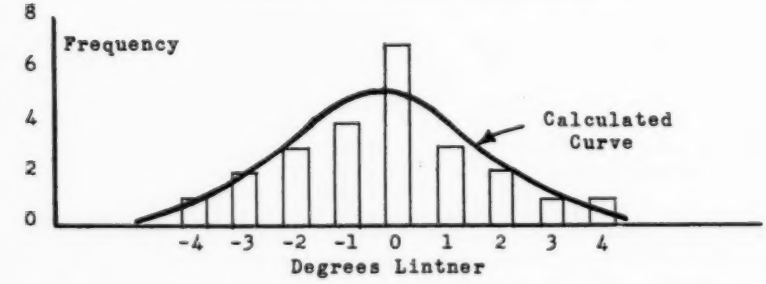
PRODUCT AND ANALYTICAL VARIATION

MALT FACTOR: DIASTATIC POWER

Frequency Distribution For Product Variation
One Malt Sample Divided For 48 Analyses



Frequency Distribution For Analytical Variation
Difference Between Duplicate Determinations
One Malt Sample Divided For 24 Analyses



sign and analysis of experiments wherein effects of using different malts can be evaluated.

REVISING SPECIFICATIONS

Up to this point we have discussed how statistics are used in considering the factors which go into a set of specifications, and this information is used when our opinion is requested in establishing a program of Statistical Quality Control. However, in many instances, the specifications are in effect and one is faced with the task of "undoing" something that has no meaning and replacing it with a sound specification. For instance, take an occasion where the consumer's specification for color was 1.5-1.7 degrees Lovibond. The first shipment of malt against this specification, comprising ten cars, was sent out with the colors within this range. The average difference between laboratories, as shown in the following tabulation, was noted when the consumer's data were received.

Wort Color Data - Degrees Lovibond

Car	Supplier	Consumer
1	1.55	1.80
2	1.55	1.75
3	1.55	1.75
4	1.55	1.80
5	1.60	1.65
6	1.55	1.75
7	1.60	1.80
8	1.55	1.65
9	1.60	1.85
10	1.60	1.70
Sum ($\sum$)	15.70	17.50
Average	1.57	1.75

Average Diff. ($\bar{D}$) = 0.18

A recheck of the results from the supplier's control data showed that their analytical method had not been out of control at the time the above shipments were made and a conference with the consumer failed to resolve the difference. The only alternative for the supplier was to establish a new specification for their own shipping department.

At this point, the human tendency is to say, subtract the 0.18°L average difference from each end of the specification and ship 1.32-1.52°L, but instead of doing this, a statistical analyses of the data was carried out to determine if:

1. The 0.18°L average difference observed for only ten samples is significant, and
2. The ranges or variabilities of the two sets of

figures, regardless of their respective levels, are the same.

The answers to these two questions are found through the use of the statistical tools called the t and F tests. Briefly reviewing, the t test is used to determine if there is a significant difference between the averages of two sets of data. In this case the 0.18°L average difference tested to be highly significant thus establishing that there was a difference between laboratories. The F test is used to compare the variances present in two sets of data regardless of their respective levels. In this example we wish to determine if the range or spread of the consumer's data is comparable to the supplier's even though the two sets average 0.18°L units apart. The results of the F test showed that the consumer's data exhibited a significantly greater variance than the supplier's.

The information obtained from the t and F tests dictated that in addition to shipping at a lower level, there could be essentially no variation or range allowed. Therefore it was requested of the production department that until we had more analytical data for comparison, an endeavor be made to ship as close to 1.45°L as possible. Figure 3, which shows a graphical history of this case, reflects the effect of this recommendation in shipments 11-20. The F and t tests were again used in analyzing the results of shipments 11-20. The difference between laboratories had dropped to 0.13°L and a decrease in variance in the consumer's data can also be observed.

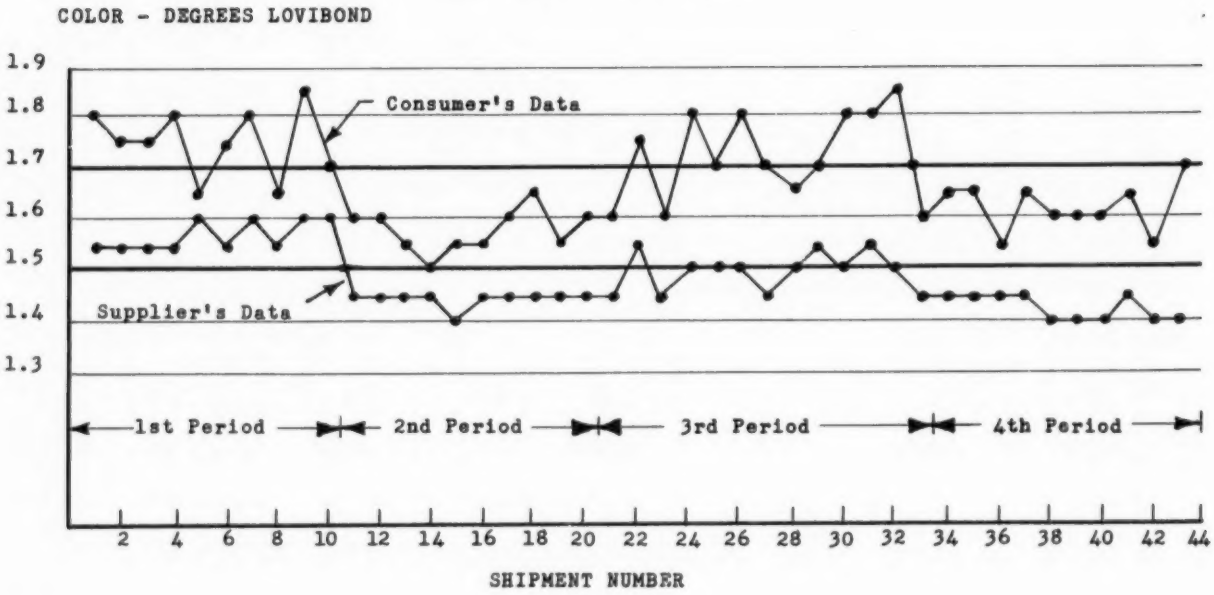
It is important at this point, to digress briefly in explaining the position of the supplier's production department, for one may not immediately appreciate all the factors involved. In the first place, this particular color requirement, 1.5-1.7°L by the consumer, was not the predominant average color desired by the industry and it meant a special product had been made. Secondly, malt processing by a supplier takes place two to four months prior to evaluation by the consumer. Where sizable volumes of malt are involved and an unusual interlaboratory difference develops unexpectedly, one can quickly appreciate the inventory accumulation problem. Therefore, in view of the improved interlaboratory situation, a series of shipments were planned in the range of 1.45-1.55°L. The results of the next 13 shipments are shown as the third period in Figure 3. Six of the cars were found out of specification by the consumer and the ratio of variances gave an F value discouragingly high. Also, the interlaboratory difference had gone up to 0.2°L.

Two simultaneous actions followed. First, in view of the increased average difference, an effort was again made to ship in the very narrow color range of 1.40-1.45°L. Second, a meeting with the quality control department of the consumer was requested to discuss the problem. Both actions brought results. The analytical data of the additional cars shipped in the color range of 1.40-1.45°L during the time the discussions were going on, are shown as the fourth period in Figure

FIGURE 3

INTERLABORATORY ANALYTICAL COMPARISON

MALT FACTOR: WORT COLOR



3. All cars met the consumer's range of 1.5-1.7°L but the extreme variance is again noted. These facts were presented to the consumer. He readily recognized that the specification coupled with the interlaboratory analytical variation were producing an unnecessary uncertainty. The color specification was, therefore, changed to 1.5-1.8°L. Subsequent shipments were scheduled in the 1.40-1.55°L range and better than 95% of the cars were then found within specification by the consumer's analyses.

A second example of how statistics were used to undo an artificial specification is shown in Figure 4. In this case, the specification was on protein ratio and called for a range of 39-41%. The histogram at the top was constructed using the analyses obtained by the supplier's laboratory for 45 shipments. The histogram at the bottom represents the consumer's analyses on the same 45 cars. Both laboratories average 40% but the difference in analytical variation is immediately apparent. It was recommended to this consumer, that because of the differences in analytical variation, the specification should be changed to 38.5-41.5%. The recommendation was accepted. The supplier continued to ship 39-41%, the consumer continues to find the values 38.5-41.5% and both are happier.

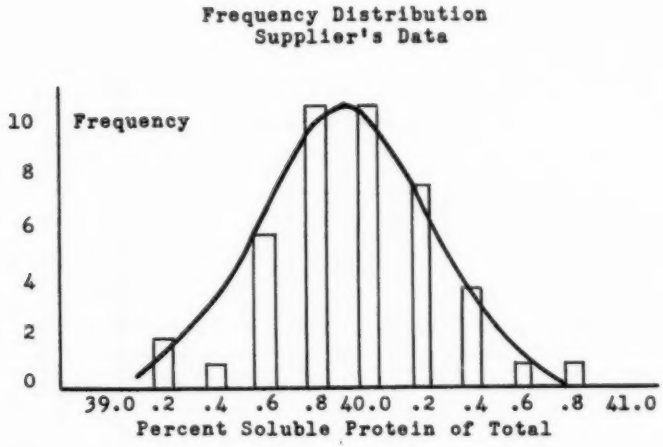
In some instances, the characteristics of the raw material dictate to some extent how the specification must be written. An example of this can be found in a study of the distribution curves for acrospire development such as shown in Figure 5. The diagram on the left shows the percentage in each growth classification for what would be considered a "short grown" malt, the one in the center a "normal grown" malt and the one on the right a "long grown" malt. The increase in percentage in the overgrown classification, as one follows from left to right, is an inherent property of barley and many other germinating seeds. Hence, a specification which calls for an extremely high percentage in the 3/4-1 growth group and is further qualified with, "overgrown-none", is ignoring "what comes naturally". In cases where specifications such as this are encountered, curves similar to the ones shown in Figure 5 are used to effect a better understanding of what is involved and the establishment of operable limits for consumer and supplier.

The very simplest of cases were used in these examples. The problem is multiplied many fold when the supplier is presented with multiple specifications, anywhere from 5 to 15 individual items. The problem is further complicated when an interdependency exists in one or more groups of specifications. A good example of this is color, diastatic power and variety. Color and diastatic power respond in opposite ways to the kilning or drying step of malting, color increasing and diastatic power decreasing. Yet it is not unusual to be asked for low color and low diastatic power and have a variety specification which restricts the use of malt having inherently low color and low diastatic power.

FIGURE 4

INTERLABORATORY ANALYTICAL COMPARISON

PROTEIN RATIO ANALYSES OF 45 SHIPMENTS



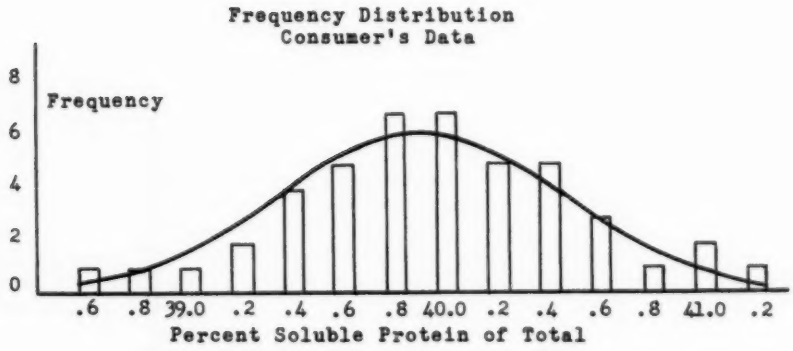


FIGURE 5

DISTRIBUTION OF ACROSPIRE DEVELOPMENT

MALTS GROWN TO VARIOUS DEGREES

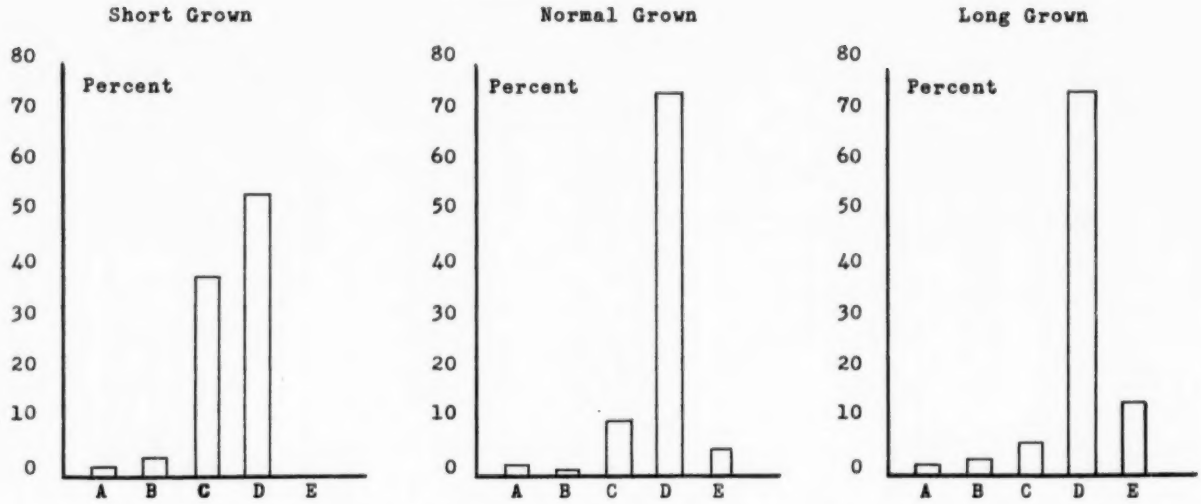


Fig. 5. Distribution of Acrospire Development; Malts Grown to Various Degrees
Vertical bars, left to right respectively in each graph: A, 0-$\frac{1}{4}$; B, $\frac{1}{4}$-$\frac{1}{2}$; C, $\frac{1}{2}$-$\frac{3}{4}$; D, $\frac{3}{4}$-1; E, over 1

CONCLUSION

There is much excellent collaborative work being carried out by committees and individual companies in our industry aimed at standardizing laboratories and establishing agreeable estimates of sample and analytical variation. The examples presented here show how a statistical treatment of similar information was used to interpret data and establish sound specifications. It is hoped, that this paper will serve to encourage others in the use of statistics and that the end product of all this work will be a better understanding and acceptance of the factors which must be considered in writing specifications.

APPLICATIONS OF SCHEFFÉ 'S METHOD OF PAIRED COMPARISONS IN FOOD QUALITY EVALUATION

O. Dykstra, Jr.
General Foods Corporation

1. Introduction

Paired comparisons techniques are being widely applied throughout industry. Of the available methods of analyzing paired comparisons experiments, we shall concern ourselves with Scheffé's analysis of variance procedure (1952). We will first review the experimental set-up and the mathematics of the Scheffé procedure. Then we will describe and illustrate some of the usual applications of the procedure.

2. Summary of Scheffé's Procedure

There are m treatments to be compared in all M=m(m-1)/2 possible pairs. We obtain 2r observations per pair, r in each order of presentation. The observations are scored statements of preference between pair members and range on a 7-point scale between +3 and -3. The preference scores are positive if the first member of a pair is preferred, negative if the second member is preferred, and 0 if there is no preference. The points of the scale are described by Scheffé as:

Score	Statement on Pair (i,j)
3	I prefer i to j strongly
2	I prefer i to j moderately
1	I prefer i to j slightly
0	No preference
-1	I prefer j to i slightly
-2	I prefer j to i moderately
-3	I prefer j to i strongly

Considerable difficulty was encountered when Scheffé's method was first applied, but by experimenting with our interviewing technique and the wording of the preference statements, we were able to find a technique which yielded nicely distributed results.

It should be noted that 5-point and 9-point scales may also be used. Again, it would be necessary to establish a proper testing technique before using these scales.

We define x_{ijk} as the k^{th} observation on the ordered pair (i,j) and assume all x_{ijk} are independent random variables with mean μ_{ij} and variance σ^2.

The mean preference for treatment i to treatment j is μ_{ij} when presented in the order (i,j) and $-\mu_{ji}$ in the order (j,i). The average preference π_{ij} and the average difference due to order of presentation δ_{ij} are given by:

$$(1) \quad \pi_{ij} = \frac{1}{2}(\mu_{ij} - \mu_{ji}), \quad \pi_{ji} = -\pi_{ij},$$
$$\delta_{ij} = \frac{1}{2}(\mu_{ij} + \mu_{ji}), \quad \mu_{ji} = \mu_{ij}.$$

The average order effect δ is given by

$$(2) \qquad \delta = \sum_{ij} \mu_{ij}/(2M),\ i \neq j,$$
$$= \sum_{i<j} \delta_{ij}/M$$

where M equals the number of pairs.

Scheffé, in his "hypothesis of subtractivity", postulates that there exist parameters $\alpha_1, \alpha_2, \ldots, \alpha_m$ characterizing the m treatments, such that the average preference π_{ij} for i to j is the difference between the corresponding parameters

$$(3) \qquad \pi_{ij} = \alpha_i - \alpha_j,\ \pi_{ii} = 0,$$

so the "deviations from subtractivity" γ_{ij} are given by

$$(4) \qquad \gamma_{ij} = \pi_{ij} - (\alpha_i - \alpha_j).$$

He adds the convenient assumption, $\sum_{i=1}^{m} \alpha_i = 0$.

The α_i may be regarded as main effects.

Since the x_{ijk} are random variables, e_{ijk} is used as the "error" in x_{ijk}, so that $e_{ijk} = x_{ijk} - \mu_{ij}$. The components which have been discussed are summarized below, together with the appropriate sums of squares resulting from partitioning the total sum of squares, using the analysis of variance table suggested by Scheffé:

Analysis of Variance

Source	Degrees of Freedom	Sum of Squares	H_o:
Main effects $(\alpha_i - \alpha_j)$	m-1	S_α	$\alpha_i = 0$
Deviations from subtractivity (γ_{ij})	M-(m-1)	S_γ	$\gamma_{ij} = 0$
Average preferences (π_{ij})	M	S_π	
Order effects (δ_{ij})	(M)	(S_π)	
Average order effect (δ)	1	$2rM\hat{\delta}^2$	$\delta = 0$
Diff. among order effects $(\delta_{ij}-\delta)$	M-1	$S_\pi - 2rM\hat{\delta}^2$	$\delta_{ij}-\delta=0$
Means (μ_{ij})	2M	S_μ	
Error (e_{ijk})	2M(r-1)	S_e	
Total (x_{ijk})	2rM	S_t	

This table also shows the identities between the sums of squares. To complete the analysis of variance we compute the mean squares and the F-ratios comparing each mean square to the error mean square. The last column has been included to show what the F-ratios are testing.

The following formulae will aid in estimating the various effects and computing the sums of squares:

Effect	Estimation	Sums of Squares
x_{ijk}		$S_t = \sum_{i,j} \sum_{k} x_{ijk}^2$
μ_{ij}	$r\hat{\mu}_{ij} = \sum_{k=1}^{r} x_{ijk}$	$S_\mu = \frac{1}{r} \sum_{i,j} (r\hat{\mu}_{ij})^2$
π_{ij}	$2r\hat{\pi}_{ij} = r\hat{\mu}_{ij} - r\hat{\mu}_{ji}$	$S_\pi = \frac{1}{2r} \sum_{i<j} (2r\hat{\pi}_{ij})^2$

Effect	Estimation	Sums of Squares
δ_{ij}	$2r\hat{\delta}_{ij} = r\hat{\mu}_{ij} + r\hat{\mu}_{ji}$	$S_\delta = \frac{1}{2r}\sum_{i<j}(2r\hat{\delta}_{ij})^2$
δ	$2rM\hat{\delta} = \sum_{i,j}\sum_{k} x_{ijk}$	$2rM\hat{\delta}^2 = \frac{1}{2rM}(\sum_{i,j}\sum_{k} x_{ijk})^2$
π_i	$2rm\hat{\pi}_i = \sum_{j=1}^{m} 2r\hat{\pi}_{ij}$	$S_\pi = \frac{1}{2rm}\sum_{i}(2rm\hat{\pi}_i)^2$

The advantage of using these formulae rather than those given by Scheffé arises from using integers throughout all the computations. The remainder of the sums of squares needed for the analysis of variance table are obtained by subtraction.

3. Illustration of the Procedure

An experiment was run in which there were 5 treatments. The 20 possible pairings, considering (i,j) as different from (j,i) were each tested 6 times. The tasters scored their preferences on the 7-point scale described earlier. The 120 preference scores are given in Table 1. These are the x_{ijk} in Scheffe's notation. The sum of the squares of these numbers give the total sum of squares in the analysis of variance table.

Also given in Table 1 are the sums of the preference scores for each pairing and represent the $6\hat{\mu}_{ij}$. The sum of the squares of these totals, when divided by 6, give the sum of squares for means in the analysis of variance table.

The $6\hat{\mu}_{ij}$ in Table 1 are used to obtain the $12\hat{\pi}_{ij}$ and $12\hat{\delta}_{ij}$ given in Table 2. For example, for the pair (1,2)

$$12\hat{\pi}_{12} = 3-5 = -2$$
$$12\hat{\delta}_{12} = 3+5 = 8 = 12\hat{\delta}_{21}.$$

For pair (1,3)

$$12\hat{\pi}_{13} = 14 - 2 = 12$$
$$12\hat{\delta}_{13} = 14 + 2 = 16 = 12\hat{\delta}_{31}.$$

As demonstrated, one table may be used to list both the $12\hat{\pi}_{ij}$ and the $12\hat{\delta}_{ij}$. In the analysis of variance table, the sums of the squares of these numbers, when divided by 12, give, respectively, the sum of squares for average preferences and the sum of squares for order effects.

The sum of all the x_{ijk}, or the sum of all the $6\hat{\mu}_{ij}$, or the sum of all the $12\hat{\delta}_{ij}$, when divided by 120, gives in the analysis of variance table the sum of squares for the average order effect.

As shown in Table 2, the $12\hat{\pi}_{ij}$ are combined to give the $60\hat{\pi}_i$. Thus:

for treatment 1, $-2 + 14 + 11 + 12 = 35 = 60\hat{\pi}_1$

for treatment 2, $-(-2) + 1 + 10 + 9 = 22 = 60\hat{\pi}_2$

for treatment 3, $-(14) - (1) + 10 + 15 = 10 = 60\hat{\pi}_3$ (Cont'd.)

for treatment 4, $-(11) - (10) - (10) - 3 = -34 = 60\,\hat{\alpha}_4$

for treatment 5, $-(12) - (9) - (15) - (-3) = -33 = 60\,\hat{\alpha}_5$.

In the analysis of variance table, the sum of squares for main effects is obtained by squaring the $60\,\hat{\alpha}_i$ and dividing their sum by 60. Note that the $60\,\hat{\alpha}_i$ sum to 0.

The remaining entries in the analysis of variance table, shown as Table 3, are obtained by subtraction. We can see that only the mean square for main effects is significant.

Thus far we have not concerned ourselves with the m treatments. They might have represented m brands of a product, as in this case. They might have levels of a multi-level factor, for which we want to determine the optimum level. Furthermore, they might have represented factorial combinations. Each of these three applications will be discussed in the following 3 sections of this paper.

4. The Treatments are Brands of a Product

We have illustrated the computations of the Scheffé procedure with an example in which 5 brands of a product were compared. We have obtained the average preference scores ($\hat{\alpha}_i$) and wish now to decide which is best, etc.

The variance between two average preference scores is

$$(5) \qquad V(\hat{\alpha}_i - \hat{\alpha}_j) = \sigma^2/rm,$$

where there are r judgments on each of the m(m-1) pairings of m treatments. In the present example $V(\hat{\alpha}_i - \hat{\alpha}_j) = .111$, so that the standard deviation of the difference between two average preference scores is .33 . We now use some multiple comparisons technique to decide which brands differ significantly in preference.

5. The Treatments are Levels of a Multi-Level Factor

We have a formulation which is fixed with respect to all but one of its ingredients. This ingredient may be the amount of sugar, the amount of color, etc. We usually have done some preliminary work and know what amount is too little and what amount is too much. The optimum which we are after will, therefore, be somewhere between these extremes. We will run the two extreme levels and one to three intermediate levels. It will be best to select the levels so that they are equally spaced in some system of measurement.

To determine the optimum we must determine what order polynomial will adequately describe the $\hat{\alpha}_i$. We set up the usual regression model and define the orthogonal contrasts for linear, quadratic, etc., as

$$(6) \qquad \hat{\theta}_j = \sum_{i=1}^{m} c_{ji}\,\hat{\alpha}_i, \quad \sum_{i=1}^{m} c_{ji} = 0;\ j=1, \ldots, m-1; \text{ and}$$

$$\sum_{i=1}^{m} c_{ji}\,c_{ki} = 0;\ j,k=1, \ldots, m-1;\ j \neq k.$$

The sum of squares for main effects S_q is partitioned into m-1 components, each with one degree of freedom, with each contrast $\hat{\Theta}_j$ contributing

$$(7) \qquad 2rm\,\hat{\Theta}_j^2 / \sum_i c_{ji}^2 = \left[\sum_i c_{ji}(2rm\,\hat{q}_i)\right]^2 / (2rm) \sum_i c_{ji}^2$$

to S_q. These sums of squares are tested against the experimental error to determine their statistical significance. Only those $\hat{\Theta}_j$ which are significant are used to estimate the $\hat{q}_i$.

The estimates of the $\hat{q}_i$ are given by

$$(8) \qquad E(\hat{q}_i) = \sum_j c_{ji}(\hat{\Theta}_j / \sum_i c_{ji}^2),$$

and the variance of these estimates is

$$(9) \qquad V[E(\hat{q}_i)] = \frac{\sigma^2}{2rm} \sum_j [c_{ji}^2 / \sum_i c_{ji}^2],$$

where the summation over j extends to only as many terms as are significant.

The optimum level is found by differentiating (8) with respect to the original variable. The derivative is set equal to zero and solved for the optimum. If the optimum is not one of the levels which we have run, we do not have an exact formula for estimating the variance of the optimum. A suggested approximation is

$$(10) \qquad V[\max E(\hat{q}_i)] = \frac{\sigma^2}{2rm} \sum_j [(c_{jo})^2 / \sum_i c_{ji}^2], \text{ where the } c_{jo} \text{ are}$$

found by inserting the level at which the optimum occurs into the formulae for the c_{ji}.

To illustrate the procedure, we have used an experiment in which there were 5 equally spaced levels of sugar. There were 18 judgments on each of the 20 pairings. In Table 4 are given the $2rm\,\hat{q}_i = 180\,\hat{q}_i$, the c_{ij} for the linear, quadratic, cubic, and quartic components (i.e., j=1,2,3, and 4, respectively). The $\hat{\Theta}_j$, their respective sums of squares, and the F-ratios based on the experimental variance of 2.83 are given. We find that only the linear and quadratic terms are significant, so that we can express the $\hat{q}_i$ as $\hat{q}_i = .6802 + .1098(i-3) - .3401(i-3)^2$. This yields a maximum for $i = 3.16$, at which level the predicted $\hat{q}_i$ is .69. The variance of the prediction is $(.0670)^2$ for i=3 and $(.0519)^2$ for i=4. Based on equation (10) we might guess that the variance of the optimum is $(.0655)^2$.

6. The Treatments are Factorial Combinations

The procedure described above may easily be extended to apply also to factorial designs. We treat the $\hat{q}_i$ as our observations as we ordinarily do, but all the resulting sums of squares must be multiplied by 2rm. If the $2rm\,\hat{q}_i$ are used as the observations, then the sums of squares must be divided by 2rm. These sums of squares are, of course, a partitioning of the main effects sums of squares in the analysis of variance table.

In the case of 2^n factorials or 2^{p-q} fractional factorials all c_{ji} equal plus or minus 1, so that the factor effects are given by

$\hat{\theta}_j/m$, which has variance

(11) $\quad \mathrm{Var}(\hat{\theta}_j/m) = \sigma^2/2rm^2$,

which is $1/(m-1)$ times the variance of the $\hat{q}_i$.

To illustrate the procedure we have selected a 2x2 factorial, in which there were 10 repetitions on each of the 12 pairings of the 4 factorial combinations. The values of $80\,\hat{q}_i$ are given in Table 5, where the factorial contrasts are also given. The F-ratios, obtained by dividing the respective sums of squares by the experimental error variance of 2.98, indicate that both the factors are showing significant effects on preference and that the interaction term is negligible.

7. Summary

We have reviewed the mathematics of the Scheffé method of paired comparisons and have illustrated the computational techniques. We have discussed and illustrated the application of the procedure when

a. the treatments are brands of a product,
b. the treatments are levels of a multi-level factor,
c. the treatments are factorial combinations.

By extending the ideas given in this paper it would be possible to apply the Scheffé procedure to the response surface type of design. This type of design is analyzed by regression analysis, and we note that applications b. and c. are relatively simple regression problems.

Reference:

Scheffé, H. (1952). An Analysis of Variance for Paired Comparisons. J.A.S.A., 47,381.

Table 1. Summary of Preference Scores

Treatment (i)	Treatment (j) 1	2	3	4	5
1		0 1 3 1 -1 -1 3	3 2 0 1 3 -1 8	0 3 1 -1 3 -1 5	-3 -1 3 3 1 0 3
2	2 -2 1 3 1 0 5		3 -2 -2 0 -3 0 -4	3 1 1 3 1 -2 7	-1 3 3 1 3 3 12
3	1 -2 -2 0 0 -3 -6	0 0 -2 0 -2 1 -3		0 1 0 -3 3 3 4	2 2 1 3 3 0 11
4	0 1 -3 -3 1 -2 -6	3 -2 -2 -1 -1 0 -3	0 -2 -2 -3 3 -2 -6		-2 -1 0 2 2 2 -1
5	-3 -1 3 -2 -3 -3 -9	1 -1 1 -1 2 1 3	-1 -3 0 3 -1 -2 -4	-2 2 1 1 1 -1 2	

Note: 1. The cells show $X_{ij1}\ X_{ij2}\ X_{ij3}$ The judges' preference scores.

$X_{ij4}\ X_{ij5}\ X_{ij6}$

$6\hat{\mu}_{ij}$ The sum of the scores for a given i and j, when i is tasted before j.

2. The preferences are scored: 3 - strong preference
2 - moderate preference
1 - slight preference
0 - no preference.

3. The preference scores are positive if the first sample tasted is preferred and negative if the second sample is preferred.

Table 2. Further Summary of Preference Scores*

Treatment (i)	Treatment (j) 1	2	3	4	5	$60\hat{\alpha}_i$	Avg.Pref. Score ($\hat{\alpha}_i$)
1		-2	14	11	12	35	.58
2	8		1	10	9	22	.37
3	2	-7		10	15	10	.17
4	-1	4	-2		-3	-34	-.57
5	-6	15	7	1		-33	-.55

* $12\hat{\pi}_{ij} = 6\hat{\mu}_{ij} - 6\hat{\mu}_{ji}$ are given in the portion above the diagonal. These are combined to yield the $60\hat{\alpha}_i$ column.

$12\hat{\delta}_{ij} = 12\hat{\delta}_{ji} = 6\hat{\mu}_{ij} + 6\hat{\mu}_{ji}$ are given in the portion below the diagonal.

Table 3. Analysis of Variance

Source	d.f.	S.S.	M.S.	F-ratio
Main Effects	4	67.57	16.89	5.06
Deviations from Subtractivity	6	14.18	2.36	.71
Average Preferences	10	81.75		
Order Effects	10*	37.42*		
Avg. Order Effect	1	3.68	3.68	1.10
Differences among order effects	9	33.74	3.75	1.12
Means	20	119.17		
Error	100	333.83	3.34	
Total	120	453		

* Not included in the totals because of partitioning of the degrees of freedom and sums of squares.

Table 4. Illustration for a Multi-Level Factor

i	$180\hat{\alpha}_i$	Coefficients C_{1i}	C_{2i}	C_{3i}	C_{4i}
1	-156	-2	2	-1	1
2	34	-1	-1	2	-4
3	110	0	-2	0	6
4	105	1	-1	-2	-4
5	-93	2	2	1	1
$\hat{\theta}_j = \sum_i c_{ji}(180\hat{\alpha}_i)$		197	-857	-79	-145
$D_j = 180 \sum_i c_{ji}^2$		1800	2520	1800	12,600
$\hat{\theta}_j^2/D_j$		21.56	291.45	3.47	1.67
F-ratio		7.62	102.99	1.90	.59

Table 5. Illustration for a 2 x 2 Factorial

i	Treatment Combination	$80\hat{\alpha}_i$	C_{ji} A-effect	B-effect	AB interaction
1	(1)	-61	-1	-1	1
2	a	12	1	-1	-1
3	b	-32	-1	1	-1
4	ab	81	1	1	1
$\hat{\theta}_j = \sum_i c_{ji}(80\hat{\alpha}_i)$			186	98	40
$D_j = 80 \sum c_{ji}^2$			320	320	320
$\hat{\theta}_j^2/D_j$			108.11	30.01	5.00
F-ratio			36.28	10.07	1.68

THE CAA PROGRAM IN QUALITY CONTROL

V.G. Moser
Civil Aeronautics Administration

Quality requirements increase in direct proportion to speed and complexity of the aircraft. Higher speeds and temperatures, much greater power in terms of horsepower and thrust, higher wing loading and faster landing speeds are taxing available materials to the limit. Each part must carry its portion of the load and do it close to its limit load. Defects obviously become less tolerable. Inspection techniques must keep step to assure sound parts.

Less than 40 years ago aircraft speeds were in the order of 100 mph. Today we consider 550 mph commonplace. Furthermore, the science of Astronautics is rapidly being developed. A pregnant question seems to be, "Is the art of quality control keeping pace with this increase of speed in flight?"

The answer to that question may very well come from the combined efforts of men in this conference. It is a well-known fact that there is a tremendous backlog of orders for civil aircraft products. It occurred to me that the subject of "The CAA Program in Quality Control" would be of particular interest to you at this time.

In the early days of aviation it became apparent to those in industry and the Congress of the United States that if civil aviation was to progress and take its place among the accepted methods of transportation, it must be safe. The Air Commerce Act of 1926 was the first federal law regulating civil aviation. This law was modernized by the Civil Aeronautics Act of 1938. This Act as amended established two civil aviation agencies with separate areas of responsibility. First, there is an agency known as The Civil Aeronautics Board, one of whose functions is to issue the Civil Air Regulations. Then there is the Civil Aeronautics Administration which has a wide variety of responsibilities in the aviation field, including such things as enforcement of safety regulations, evaluating new aircraft from a safety standpoint, licensing airmen and helping municipalities improve their airports.

Enforcement of safety regulations and evaluating new aircraft from a design and manufacturing standpoint puts the Civil Aeronautics Administration in the field of quality control. Quality Control is synonymous with safety, and safety is CAA's prime objective in the program for the promotion and advancement of civil aviation. As you know, safety does not "just happen". We in CAA recognize and appreciate the valuable contribution of groups such as this which have advanced safety in civil aviation to where it is today.

For the purpose of this discussion we will concentrate on CAA responsibilities as they apply to the design, construction, quality control, and certification of aircraft and related products. From the time a new aircraft is on the drawing board CAA, through its office of Flight Operations and Airworthiness, works to assure its safety. First comes cooperative effort with the manufacturers to see that an ample margin of safety is built into the vehicle itself.

There are three basic certificates which CAA issues to signify

that industry has met the pertinent requirements. These will be discussed in chronological order.

First, there is the type certificate, which is a document indicating that the design of the product meets the requirements specified by the Civil Air Regulation. In order to obtain a type certificate, the manufacturer must first submit data to CAA that defines and discloses the configuration and performance of the airplane, engine or propeller. This includes drawings, specifications and test reports.

Second, a production certificate is a document issued to a manufacturer when he has demonstrated to the CAA that he is consistently capable of producing articles with adequate quality control conforming to the approved type design. This involves an examination of the manufacturer's facilities by a group of specialists. They will determine that the manufacturer has adequate written procedures on all the salient facets of production, trained personnel and equipment to assure production that meets required quality and safety standards.

Third, an airworthiness certificate is issued for an aircraft as evidence that it conforms to CAA approved type design and that it is safe for operation. Each CAA approved aircraft is issued an airworthiness certificate which remains in effect as long as the aircraft is maintained in accordance with the Civil Air Regulations.

The regulations that CAA administers are based on facts gained through service experience over the past years. They are based on facts, not opinions. The aviation industry has provided many of the facts around which regulations have developed. Industry has contributed in a large measure to the drafting of these regulations. Before new regulations are adopted they are submitted in the form of a proposal to the aviation industry for comments on whether compliance is economically feasible and whether in their opinion the safety objectives proposed will be attained. Joint meetings are held in which the CAA, the CAB and industry representatives all have a part in drafting of the final regulation. From this viewpoint then the CAA is an enforcement agency enforcing Civil Air Regulations which have largely been developed in cooperation with the industry.

The Aircraft Engineering Division is a segment of the CAA whose responsibility it is to evaluate new civil aeronautical products for compliance with regulations and for eligibility of the product for CAA certification. Safety, airworthiness, and reliability are the prime considerations for certifications. On the staff of the Aircraft Engineering Division are such specialized personnel as structural engineers, aerodynamicists, vibration specialists, power plant engineers, propeller specialists, test pilots, flight analysts, and manufacturing quality control specialists.

It is their responsibility to evaluate the product presented by the manufacturer to determine whether it meets the requirements of the Civil Air Regulations. This staff of specialists also evaluates the testing programs proposed by the manufacturers in connection with new designs, new materials, and new processes. Processes and procedures are carefully evaluated in factories, and it is in this evaluation that quality control systems come under CAA examination. There is a significant difference between quality control from the viewpoint of the manu-

facturer and quality control as it concerns CAA. In addition to safety, the manufacturer's quality control is concerned with such things as customer acceptance, price, eye appeal, etc. The CAA is concerned only with those quality characteristics that affect safety. It is a CAA responsibility to ascertain that production and quality control systems are adequate to produce duplicates of articles conforming to the approved type design.

The CAA has endeavored to develop a small efficient staff of quality specialists to cover all manufacturers producing civil aircraft products such as aircraft, engines and propellers. It is obvious that with a small group of people we cannot expect to conduct a detailed inspection of all products involved. Therefore, we must depend on our system of evaluation of a manufacturer's procedures and quality control system.

A great deal of thought has been devoted to the problem of how we can utilize our personnel most efficiently in accomplishing our responsibilities for verifying continued compliance on the part of the manufacturer. It has been found that the facilities of any company producing aircraft products can be subdivided into functional areas. These areas are then evaluated periodically and systematically by the use of survey reports, control charts and follow-up. In conducting these evaluations, we make a determination of six items which are common to every type of area, regardless of function. These items include: personnel, facilities, technical data, product conformity, records, and general compliance. Our experience has shown that the use of this surveillance system will furnish us the most accurate evaluation of the manufacturer's operation with a minimum of man-power. We do not duplicate the manufacturer's quality control system, but verify the results of his system.

Service records provide a barometer and additional tool for evaluating the quality of a manufacturer's product. Service difficulty information comes from many sources including airline operators, accident reports, and through our system of malfunctioning and defects reports. Malfunctioning and defects reports are submitted by aircraft owners, operators, mechanics, and by our field personnel. This information is assembled by our Washington office where it is analyzed for statistical purposes and appropriate action.

As the industry has grown and demonstrated its ability and willingness to assume more responsibility in the field of safety regulations, the CAA in turn has delegated additional authority to industry as provided for by the Civil Aeronautics Act. In this connection, qualified industry personnel may be authorized to perform certain functions on behalf of the CAA, under CAA supervision. In so doing they may be authorized to approve aeronautical products including airworthiness certification of aircraft under the standards established in the Civil Air Regulations. This provides certain benefits to both the manufacturer and CAA in that around-the-clock service will be available. This releases CAA personnel from these duties which will permit them to more effectively perform other functions.

At this point it should be emphasized that CAA has not relinquished its responsibility for assuring safety.

Industry personnel appointed to perform CAA functions must first meet rigid qualification standards. After they are selected they are given on-the-job training and furnished with all necessary regulations and operating instructions. Their performance is subject to continuous evaluation by CAA personnel.

In order to give you an idea of the magnitude and progress of the civil aviation industry, I will quote a few statistics compiled by CAA.

	1955	1956	Percentage Increase
Registered civil aircraft Manufactured	4820	7205	49.5%
Civil airframe weight	10,230,500	16,055,900	56.9%
Civil engines	7639	11,501	50.6%
Civil engine (h.p.)	3,338,400	5,656,600	69.5%

On June 30, 1956 there was a total of 83,841 civil aircraft registered in the U. S.; 63,092 were what we call active aircraft.

As you can see from these statistics, the importance of the civil aircraft industry has reached major proportions. I believe that improved quality control methods, plus the increased interest in quality control by management, has played an all important roll in the amazing progress of the industry.

The Civil Aeronautics Act was designed not only to assure safety in air transportation, but also to promote aeronautical development and experimentation. For this reason the Civil Air Regulations are written objectively. Since the CAA is not the customer, the manufacturer is given wide latitude in the design and manufacture of his products, provided the minimum standards of safety are met.

It is a fundamental policy of CAA to encourage industry to develop effective methods of quality control which will assure continued growth of the industry through improved safety.

It is my opinion we too often consider quality control as a non-productive facet of production. The aircraft industry is a mature industry, and is the largest industrial employer in the U. S. I am sure you will agree experience has demonstrated that quality is the cornerstone for a dynamic and safe civil aircraft industry.

VARIABILITY OF MECHANICAL PROPERTIES OF FLAT ROLLED SHEET PRODUCT

John V. Sturtevant
U. S. Steel Corporation

Variability of mechanical properties of flat rolled metal product is a matter of keen interest, not only to its fabricators, but equally to its producers who are constantly working to improve the uniformity of these properties. Under the impetus of increased scientific knowledge of the nature and behavior of metals, our modern equipment, technology, and instrumentation have been combined by alert management to provide degrees of uniformity which are markedly superior to those previously attained. These efforts continue, for it is well recognized that the fabricators' interests are best served by products having mechanical properties whose variabilities are at the minimum levels consistent with sound economic practices. Contributions of the technical and production staffs of the fabricators' plants must be equally recognized, for without their enthusiastic and whole-hearted cooperation, much of the progress which has been made to date could not have been accomplished. This splendid spirit of cooperation is most heartening and, as a member of the Metals Technical Committee of ASQC, I welcome this opportunity to make some small contribution toward a better appreciation of the variabilities found in the mechanical properties of flat rolled metal product.

Before proceeding with the details of this study, I wish to explain that this paper is one of a series which is being sponsored from time to time by the Metals Technical Committee with the aim of improving our mutual understanding of this general problem of variability in mechanical properties. In view of the broad scope of this field, however, it appears preferable to consider only one property and product at a time. Accordingly, at the Montreal Convention in 1956, Dr. John W. W. Sullivan presented a similar paper concerned with the Rockwell hardnesses of steel sheets.[1] Although our present study is again concerned with a steel product, I wish to emphasize that the Metals Technical Committee is keenly aware of the existence of similar problems in the nonferrous field and hopes, in the future, to bring these before you.

Our present study concerns the variabilities found in Olsen ductility cup tests made on 25 sheets of rimmed steel taken from the middle of the lengths of 25 coils rolled straightaway on a conventional mill. It is important that we clearly recognize that the uniformity found among these test values is substantially greater than that which can generally be expected of such product since, first, the specimen sheets represent only one position with respect to coil length, second, the slabs from which these coils were rolled were substantially similar in chemical composition and third, the coils having been rolled at the same time, were subject to the same processing conditions.

In making the Olsen ductility test, a small specimen is cut from the sheet and held in position over a die while a steel ball is steadily pressed into the specimen, forming a small cup. The test value is then recorded in mils as the depth to which the ball has been forced when incipient fracture is indicated. Since we are concerned in this study only with the variability found among these test values, it has been

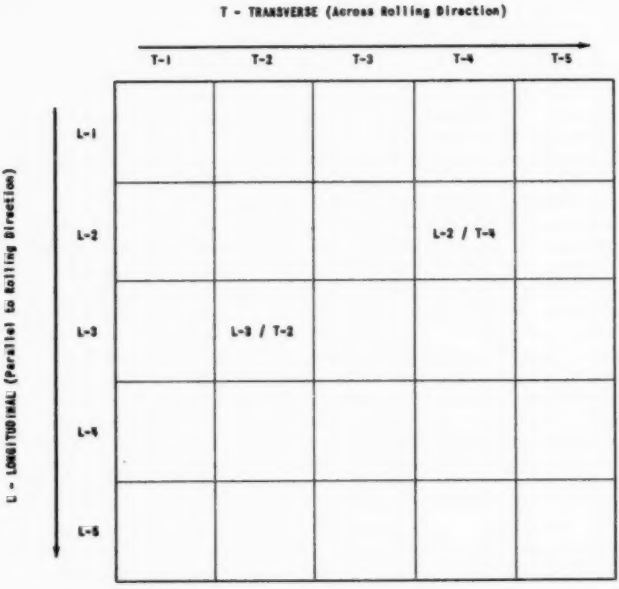


FIGURE 1 - Schematic Lay-out of Specimen Sheets Used For Ductility Cup Tests

Table 1

Percentage Distributions of Individual Values Arranged by (1) Position Along Lengths of Sheets and (2) Across Widths of Sheets

Positions Along Sheet Length						Positions Across Sheet Width					All
L-1	L-2	L-3	L-4	L-5	Mils	T-1	T-2	T-3	T-4	T-5	Values
0.8	1.6	4.0	1.6	1.6	< 1			4.8	1.6	3.2	1.92
0.8	5.6	3.2	4.0	3.2	1- 5	2.4	3.2	2.4	4.0	4.8	3.36
9.6	8.8	8.0	5.6	5.6	6-10	4.8	8.8	12.0	10.4	1.6	7.52
11.2	16.0	15.2	11.2	10.4	< 11	7.2	12.0	19.2	16.0	9.6	12.80
20.0	16.0	21.6	18.4	15.2	11-15	12.0	25.6	24.8	20.0	8.8	18.24
20.0	21.6	19.2	20.8	23.2	16-20	16.8	16.8	30.4	20.8	20.0	20.96
16.8	21.6	20.0	19.2	24.0	21-25	21.6	24.0	12.8	18.4	24.8	20.32
19.2	13.6	13.6	14.4	16.8	26-30	24.0	12.0	8.0	12.8	20.8	15.52
76.0	72.8	74.4	72.8	79.2	11/30	74.4	78.4	76.0	72.0	74.4	75.04
12.8	5.6	6.4	14.4	8.0	31-35	13.6	8.0	4.0	9.6	12.0	9.44
	4.8	4.0	1.6	1.6	36-40	4.8	1.6	0.8	0.8	4.0	2.40
	0.8			0.8	> 40				1.6		0.32
12.8	11.2	10.4	16.0	10.4	> 30	18.4	9.6	4.8	12.0	16.0	12.16

possible to simplify the presentation by considering, for each test specimen, only the number of mils by which the actual test value exceeds a constant reference value. Thus, these "coded" values range from (minus) -20 through plus 42 mils, i.e., through an over-all range of 63 mils.

Turning to Figure 1, we find the schematic layout for the preparation of the 625 individual Olsen ductility specimens. It is noted that each sheet was sheared to a length equal to its width and then sheared into five intersecting rows and columns. The rows, designated as L-1, L-2, etc., progress in the direction of rolling while the columns, T-1, T-2, etc., progress from left-to-right across the width of the sheet. As a result, we can consider the five test values representing each of the five rows and five columns as a group of size $n = 5$. Further, on passing along the series of 25 sheets, we also have a series of 25 L-1 groups, 25 L-2 groups, etc., and a similar series for the T-groups. In addition, each of the 25 sheet positions (such as L-3/T-2) can be considered as providing one of a series of 25 test values respectively representing the 25 coils. Thus, it is evident that we have several bases upon which to consider variabilities, not only among the individual Olsen ductility test values, but also among their group averages, ranges, and standard deviations.

Turning to Table 1, let us first note the per cent frequency distributions of the individual values as arranged in accordance with the five row positions (L_i) and the five column positions (T_i). Here we note in the fourth line from the top that, while the frequencies of values below 11 mils for the L-2 and L-5 rows differ by 5.6%, the T-1 and T-3 columns differ by 12.0%. At this point, let us turn to Dr. John W. W. Sullivan's paper on Rockwell hardness given at the 1956 Convention.[1] With his kind permission, I quote in part: "Rimmed steels are characterized by marked differences in chemical composition across the section and from top to bottom of the ingot. They have an outer rim that is lower in carbon, phosphorus, and sulphur than the average composition of the whole ingot, and an inner portion or core that is higher than the average in those elements."----"The structural pattern of the rimmed steel ingot persists through the rolling process to the final product." Thus we see that the centers of the widths of steel sheets taken from coils rolled in a direction parallel to the heights of their parent ingots must be expected to have carbon, phosphorus, and sulphur contents higher than those found near the edges of the sheets. Since increased contents of these elements tend to reduce ductility, the relatively higher frequency of the lower Olsen ductility values in the T-3 columns (Table 1) is to be expected. Thus we see that, on passing from edge to center of the sheet, an inherent technical characteristic of this widely-used type of steel constitutes a source of variability in the ductility test values. Upon passing from row-to-row along the length of the sheet, however, a lesser variability is to be expected since the total length of a single sheet constitutes but a very small increment of the total height of the parent ingot as poured.

In view of the differences noted in Table 1 among the frequencies of "low" (below 11 mils) and "high" (above 30 mils) test values found in the various "T" columns, Table 2 was prepared in order to exhibit these frequencies as they occur from coil-to-coil. In order to simplify the presentation, these frequencies are presented by count rather than in the form of percentages. Here we find that, despite their similarity in

Table 2

Frequency by Count of Low* and High** Values Arranged by Position Across Widths of Sheets and by Coil

Under 11 Mils*						Coil	Over 30 Mils**					
T-1	T-2	T-3	T-4	T-5	T-All	No.	T-1	T-2	T-3	T-4	T-5	T-All
						1				2	1	3
1	1			1	3	2			1		2	3
						3					2	2
						4						
						5	1			1		2
	1	1	1	1	4	6					1	1
3	1	4	4	4	16	7						
	1	1	1	2	5	8	2					2
		1			1	9	1	1				2
	2	2			4	10						
	4	3	4	1	12	11						
	3	2		1	6	12						
						13						
2	1	1	1		5	14						
		3	1		4	15					2	2
		1			1	16	1					1
						17	2	2	1		2	7
2			3		5	18						
			1		1	19	1					1
		1	1	2	4	20						
						21	1			1		2
						22	5	3	2	4	3	17
1	1	3	3		8	23						
						24	5	3	1	4	4	17
		1			1	25	4	3	1	3	3	14
9	15	24	20	12	80		23	12	6	15	20	76

Table 3

Frequencies by Count of 5-Value Averages Arranged by (1) Position Along Lengths of Sheets and (2) Across Widths of Sheets

Position Along Sheet Length							Position Across Sheet Width					
L-1	L-2	L-3	L-4	L-5	L-All	Mils	T-1	T-2	T-3	T-4	T-5	T-All
		1			1	< 1						
		1	1		2	1- 5			2	1	1	4
1	3	1	1	1	7	6-10	1	1	2	1	1	6
1	3	3	2	1	10	< 11	1	1	4	2	2	10
6	6	3	2	2	19	11-15	3	7	8	6	2	26
3	3	7	11	9	33	16-20	3	4	8	6	6	27
9	8	9	4	9	39	21-25	7	9	3	7	5	31
4	2		3	2	11	26-30	8	2	2	2	8	22
22	19	19	20	22	102	11/30	21	22	21	21	21	106
2	3	3	3	2	13	31-35	3	2		2	2	9
						36-40						
						> 40						
2	3	3	3	2	13	> 30	3	2		2	2	9

Table 4

Frequencies by Count of 5-Value Ranges Arranged by (1) Position Along Lengths of Sheets and (2) Across Widths of Sheets

Position Along Sheet Length							Position Across Sheet Width					
L-1	L-2	L-3	L-4	L-5	L-All	Mils	T-1	T-2	T-3	T-4	T-5	T-All
1	4		2		7	1- 5	4	1	2	2		9
7	5	10	15	7	44	6-10	6	13	13	6	11	49
8	9	10	17	7	51	< 11	10	14	15	8	11	58
8	11	9	2	11	41	11-15	13	6	4	6	5	34
5	1	4	3	6	19	16-20	1	3	2	10	3	19
3	2		2	1	8	21-25	1	2		1	3	7
1	1		1		3	26-30			2		2	4
17	15	13	8	18	71	11/30	15	11	8	17	13	64
		1			1	31-35			1			1
	1	1			2	36-40			1		1	2
	1	2			3	> 30			2		1	3

Table 5

Frequency by Count of Individual Values Arranged by Row/Column (L_1/T_1) Positions on Sheet

Row L	Mils	Column T-1	T-2	T-3	T-4	T-5	T-All
	< 11	1	3	6	3	1	14
1	11/30	20	21	17	19	18	95
	> 30	4	1	2	3	6	16
	< 11	2	3	5	5	5	20
2	11/30	18	19	20	18	16	91
	> 30	5	3	0	2	4	14
	< 11	3	5	5	4	2	19
3	11/30	18	18	19	18	20	93
	< 30	4	2	1	3	3	13
	< 11	1	1	4	5	3	14
4	11/30	20	20	18	16	17	91
	> 30	4	4	3	4	5	20
	< 11	2	3	4	3	1	13
5	11/30	17	20	21	19	22	99
	> 30	6	2	0	3	2	13
	< 11	9	15	24	20	12	80
All	11/30	93	98	95	90	93	469
	> 30	23	12	6	15	20	76

Table 6

Frequencies by Count of Averages of Groups of Twenty-five Values Representing (1) Entire Sheets and (2) Row-Column Sheet Positions (L_1/T_1)

Mils	Entire Sheets	Row-Column Positions
< 11	2	0
11/15	3	2
16/20	8	13
21/25	9	10
26/30	1	0
> 30	2	0

chemical composition, Coils No. 10 through 14, 18, 20, and 23 display no test values above 30 mils while Coils No. 22, 24, and 25 show a preponderance of such values. Thus, it is clearly evident that even with the diligent care which is exercised by the technical and production staffs of the producing mills, variability in ductility test values must be expected on passing from coil-to-coil.

When variability is encountered among individual data, it is, of course, customary to consider the possibility of using averages of groups to represent more adequately the characteristic being studied. With this aim, the averages of the groups of 5 values representing the rows and columns on the 25 specimen sheets are listed in Table 3 in the form of frequencies by count. While the differences are not outstanding, the T-3 column should be compared with the other "T" columns with respect to its frequencies of test value averages below 11 mils and above 30 mils. Here we find that the stabilizing effect of the averaging process yet reveals the tendency of the center of the sheet width (T-3 column) to exhibit somewhat lower ductility values. Turning to the question of the ranges of these same groups of 5 values, it is interesting to note in Table 4 that column T-3 displays some tendency toward both "low" and "high" ranges as compared with the other "T" columns. The technical significance of this double tendency is not, however, clear.

Returning to consideration of the individual values, Table 5 lists their frequencies by count as arranged by their row/column positions (L_1/T_1) on the 25 specimen sheets. Here it is evident that there is no single position which displays an outstanding frequency of "low" or "high" test values. Instead, as previously indicated in Table 1, the effect of position across the sheet width (T-columns) strongly overshadows that of position along the sheet length.

Table 6 presents the frequencies by count of averages of groups of 25 values arranged in two manners, those representing entire sheets without regard to the rows and columns of individual test specimens and those comprising row/column sheet positions (L_1/T_1). A comparison of the two significant columns listed in this table clearly indicates that greater variability is to be expected among the sheets taken from different coils than among the row/column positions used in the test specimen layout.

The use of statistical control chart techniques is always an intriguing possibility in working with data of this nature. Let us, then, examine this set of 625 test values with respect to their statistical uniformity in the sense of Shewhart. Considering first the 10 row and column groups of size n = 5, we find, in Table 7, the conventional control chart statistics for each of the "L" and "T" groups. Turning next to Table 8, we find listed the several group averages which fall near or beyond the 3-sigma limits given in Table 7. Since these averages are arranged with respect to their coil numbers, it is interesting to note that nearly all of the averages representing Coils No. 11, 22, 23, 24, and 25 lie beyond either the lower or upper limits. It would, of course, be tempting to exclude these groups and recompute the control chart limits, but to do so would not be consistent with sound statistical principles for we are not in a position to designate their assignable causes and, further, there would yet remain 11 coils displaying averages as yet unexplained.

Table 7

Control Chart Statistics for Groups of Five Values Taken (1) Across Widths and (2) Along Lengths of Sheets - Expressed in Mils

(* - 3 - Sigma) Statistic	Statistics Computed: Across Sheet Width L-1	L-2	L-3	L-4	L-5	Along Sheet Length T-1	T-2	T-3	T-4	T-5
$\overline{\overline{X}}$ -	20.7	19.8	19.0	20.7	21.0	23.4	19.9	16.5	19.5	22.0
$\overline{R}$ -	13.5	13.6	14.0	11.7	13.6	11.2	11.5	13.7	13.7	14.8
$\overline{X}$ - LCL* -	12.9	12.0	10.9	13.9	13.1	16.9	13.2	8.8	11.6	13.4
- UCL* -	28.5	27.6	27.1	27.4	28.8	29.8	26.4	24.0	27.4	30.5
$\overline{R}$ - UCL* -	28.6	28.7	29.7	24.7	28.8	23.7	24.3	27.8	29.0	31.2
Individuals										
- LCL* -	3.3	2.3	0.9	5.6	3.4	8.9	5.0	-1.2	1.8	2.9
- UCL* -	38.2	37.3	37.1	35.7	38.5	37.8	34.7	34.1	37.2	41.0

Table 8

Occurrences Near or Beyond 3-Sigma Control Chart Limits - Averages of 5 Values Taken (1) Across Widths and (2) Along Lengths of Sheets - Expressed in Mils

Coil No.	Across Sheet Width L-1	L-2	L-3	L-4	L-5	Along Sheet Length T-1	T-2	T-3	T-4	T-5
1		31.2						23.8	28.6	
2		10.6		29.0						
3				27.4						
4										
5				29.0						
6					6.0					
7		9.6	-1.0	5.0		9.0	13.0	4.2	4.6	3.0
8										
9										
10										
11		8.4	5.0	11.0	10.6	16.6	8.4	3.0	10.0	11.0
12							12.0			
13										
14		8.4				14.2				
15										
16										
17	30.0									
18						13.6			11.8	
19										
20	9.6			10.0					11.4	10.2
21										
22	30.6	29.4	32.6	33.0	34.6	32.8	32.6	28.0	34.0	32.8
23	13.0	12.0	10.0	14.0		13.0	13.2		10.6	
24	32.4	35.4	32.8	33.0		34.6	31.6	29.0	32.8	33.4
25		31.6	32.8	31.4	31.0	32.4	28.0		28.4	

Table 9

Control Chart Statistics for Groups of Twenty-five Values Representing (1) Entire Sheets and (2) Row-Column Sheet Positions - Expressed in Mils

Category	$\bar{\bar{X}}$	$\bar{\sigma}$	3 - Sigma Limits for Averages LCL	Averages UCL	Std. Devs. LCL	Std. Devs. UCL	Individuals LCL	Individuals UCL
Sheets	20.2	6.18	16.4	24.1	3.49	8.87	1.1	39.4
Row-Column Positions (i.e., L_1/T_1)	20.2	8.17	15.2	25.3	4.62	11.72	-5.0	45.5

Table 10

Occurrences Near or Beyond 3 - Sigma Control Chart Limits - Groups of 25 Values Representing (1) Entire Sheets and (2) Row-Column Sheet Positions

Entire Sheets			Sheet Positions			
Coil No.	Average	Standard Deviation	Row-Column No. L	T	Average	Standard Deviation
1	24.3		2	5		13.1
5	25.0		3	3	14.6	12.0
7	6.8					
11	9.8	10.6				
14	14.2					
20	13.7					
22	32.0					
23	12.7					
24	32.3					
25	28.4					

Continuing to examine the possibility of using control charts, let us note the limits listed in Table 9 for groups of 25 values as arranged to represent, first, entire sheets and, second, row/column positions (L_1/T_1). In each case we are dealing with 25 averages. Owing to the larger group size, sigma-bar instead of R-bar has, of course, been computed and used in determining the control limits. Again, in Table 10, as in Table 8, the occurrences beyond the control limits have been listed. It is evident that the over-all sheet averages are lacking in statistical uniformity. The fact that the row/column sheet positions show only 2 of their 25 averages beyond the control limits tends to support our previous conclusion from Table 5 that the individual sheet positions tend to provide similar results. It appears, therefore, as unlikely that the statistical control chart technique can be appropriately used in establishing ranges of expectation for Olsen ductility test values.

Summarizing, we find, first, that Olsen ductility cup test values tend to be lower near the center of the width of sheets made from rimmed steel than near the edges of such sheets and that a source of this variability is the tendency of the molten steel in the ingot to chemically segregate toward its center. Second, the ductility values vary, but less markedly, along the length of the sheet. Third, that although diligent care is exercised in the selection and processing of steel sheets, their ductility values vary significantly on passing from coil-to-coil. Fourth, the averaging of ductility test values does not provide wholly uniform data on passing from the edge to the center of the sheet width. Fifth, the lack of statistical uniformity in the sense of Shewhart precludes the sound use of statistical control chart techniques in establishing ranges of expectation for Olsen ductility cup test values.

Reference

1. "Non-Random Distribution of Variables - Rockwell Hardnesses of Hot Rolled Steel Sheets in Coil Form", by Dr. John W. W. Sullivan, American Iron and Steel Institute, National Convention Transactions, American Society for Quality Control, 1956, pages 149 - 161

DESIGNING THE MEASUREMENT AND ANALYSIS OF WORK

John M. Allderige
Eastman Kodak Company

A large part of Industrial Engineering effort is devoted to "Work Measurement." This is put in quotes here since the activity as practiced is more than just "measuring" the time associated with some particular work activity. In practice the engineer also selects appropriate characteristics of the operation to be time-measured, he improves methods and procedures where possible, and he sets up a control system for scheduling and, in the case of incentive standards, pay whereby countable output can be related to time input. In other words, "Work Measurement" really means a complete study and description of the operation.

This analytical process has two goals, both mentioned above -

1. Improvement of the operation - using less time, with less fatigue, more safety and generally more satisfaction.
2. Control of the operation - knowing the labor content such that future labor requirements can be predicted from estimates of future production requirements.

The first goal - higher productivity and greater humanization - is usually sought through work simplification, methods engineering, economic motivation of incentive standards, and other types of motivation. The second goal - control - is sought through a quantitative description of the operation relating production and labor time. This control is desirable since it usually means better utilization of the individual operation with regards to the entire plant - that is, control usually results in higher overall improvement.

Mathematics and statistics are fundamental to the effective realization of these goals. In Statistical Quality Control, Experimental Design, and Operations Research, very powerful techniques have been developed that can be used in "Work Measurement." This paper proposes to discuss some of these mathematical and statistical techniques and their role in designing the measurement and analysis of work. Primary attention is paid to the second goal - seeking a control system - although some of these techniques can help in selecting better methods and better procedures for improving the individual operation. The area of motivation - such as incentive standards - is not within the scope of this discussion. Mathematical and statistical techniques have not yet been developed - if, indeed, conceived - for characterizing the motivation of human behavior in industrial production operations.

In seeking a control system for an industrial operation, there are four factors of primary importance: (1) The decision tolerance; (2) the decision period; (3) the mathematical model of the operation; (4) the measurement process. These four factors interact with each other and with outside factors in a fashion that can be best described by a small example. Consider a simple repetitive operation of assembling a bolt and nut. Now, the time to do 100 such assemblies would be arrived at by multiplication:

Average time per assembly x 100 = total time for 100 assemblies
say, .2 minutes x 100 = 20 minutes

This is a trivial calculation, of course, but it can be expressed mathematically with the following substitutions:

$\bar{\bar{X}}$ = average time = .2 minutes
N = production of assemblies = 100
T = total time

$$T = N\bar{\bar{X}} = (.2) \times (100) = 20 \text{ minutes}$$

where $T = N\bar{\bar{X}}$ is the mathematical model for the operation.

Now suppose the time to do an assembly, X, is subject to quite a bit of variation due to, say, discarding defective bolts, as well as to chance operator variation in performing the assembly. If we wished to come to a decision as to the labor content for every 100 assemblies we might run into trouble using $\bar{\bar{X}}$ - the average time per assembly. Our estimates might vary from reality by more than some acceptable tolerance. We would continuously estimate 20 minutes when the actual times to do successive groups of 100 assemblies might vary from, say, 15 to 25 minutes. (This can be treated statistically, of course, but that will come later. The point here is to identify the problem in general terms.) The greater the variation in time per assembly, the greater the variation in the time for 100 assemblies.

If we wanted to make an estimate that was not more than 1 minute off in each direction, we would then have to analyze the operation further to see what was causing the variation. One suggestion as to an assignable cause is the handling of defective bolts. That is, some assemblies require picking up and discarding a defective bolt before making a good assembly. Suppose we decided to count such defective bolts in the future and associate a time with the handling of defective bolts. This is referred to as "selecting another unit of measure" but, more fundamentally, it is changing the mathematical model. Our calculation now for the time to do 100 assemblies is:

Time for 100 assemblies =

time for 1 good assembly x 100 assemblies + time for handling a defective bolt x number of defective bolts for the 100 good assemblies.

If $\bar{\bar{X}}_1$ = average time for 1 good assembly

N_1 = number of completed assemblies = 100 in this case

$\bar{\bar{X}}_2$ = average time for handling a defective bolt

N_2 = number of defective bolts

T = total time

Then the mathematical model becomes:

$$T = \bar{\bar{X}}_1 N_1 + \bar{\bar{X}}_2 N_2$$

Whether or not successive T's - estimates of time for the 100 assemblies - would more closely match actual T's than before would be a measure of this model's success. If it still was inadequate - if the residual variation was still too high - then more study would be required to develop a more extensive mathematical model.

We have just demonstrated one interacting effect - that of the mathematical model on the decision tolerance. Given a decision tolerance and a decision period - in this case within 1 minute over 100 assemblies - the engineer can seek a sufficiently descriptive mathematical model for his control system by being aware of the variability in the system. This is essentially a statistical problem as will be reviewed shortly. Note, however, that it may not have been possible to expand the model. Supposing in making an assembly variability existed due only to chance operator causes such as widely differing methods. From the point of view of devising a control system, on the existing operation, these method differences are not countable and the engineer would have to conclude that either the decision tolerance and/or the decision period needed modification. Something would have to give.

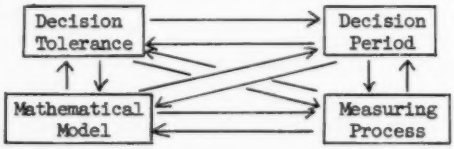
Widening the decision tolerance, of course, is an obvious device - if the system can only give you a 2 minute tolerance and you now decide to accept 2 minutes instead of 1 minute, then the problem is resolved. Lengthening the decision period may not be so obvious to those unfamiliar with statistics but it is reasonable when studied. By shifting from 100 to 200 assemblies as the decision period one would reduce the effect of the individual assembly variation. That is, relative to some average time for 200 assemblies, the actual times for successive groups of 200 assemblies vary less than that same process with 100 assemblies. Put in even different words - the average times for groups of 200 assemblies bounce around less than the average times for groups of 100 assemblies.

The above demonstrates the interaction between the mathematical model, the decision tolerance, and the decision period. The latter two are not always as inviolate as their name implies although they often exist as policy. In incentive standards situations, for instance, it may be policy (tradition) to post earnings (percent effectiveness) daily and to require that there be no more than (say a 5% risk of no more than) plus-or-minus 3% variation for reasons beyond operator control (such as the occurrence of defective bolts mentioned above). Lengthening this decision period to one week might not be impossible if pay checks are issued weekly. Similarly in production programming arriving at an appropriate decision period depends on such factors as the mobility of the labor force, the degree of refinement required in meeting deadlines, and the degree of refinement required in cost estimating. It isn't the purpose of this paper to explore the specification of the decision period and the decision tolerance but to point out the relationship of the operation itself and these constraints.

One other possibility in this triangle exists - changing the operation itself. Work simplification, methods engineering, training, and the improvement of product quality into the operation might all be possible ways of reducing methods differences. Of course, one traditional way is to use the time associated with a particularly good method as a "standard" and as a base for premium production pay. This and the other approaches are only successful from a control point of view if they achieve reduced variation and, therefore, better control. Incidentally, in such situations - wide unexplainable variation - pay control for

incentive standards can be brought within tolerance by a fixed allowance.

So, we have described three of the four factors and some of the ways they interact. The decision tolerance and the decision period may require a mathematical model impossible to obtain with the chance variation in the operation. Changing the decision period and/or the decision tolerance and/or the operation can bring things into a compatible state. The fourth factor, the measurement process, has to do with the engineer's activities. In our bolt and nut example, for instance, having decided to test the mathematical model $T = \bar{\bar{X}}_1 N_1 + \bar{\bar{X}}_2 N_2$ we would then be required to obtain, by time measurement, $\bar{\bar{X}}_1$ and $\bar{\bar{X}}_2$. Suppose X_2, the time to handle a defective bolt, varied greatly and therefore required an enormous number of observations for adequate precision. This data collection cost might well cause some second thoughts regarding the use of this expanded model and/or suggest reappraisal of the decision period and/or the decision tolerance. So, the measurement process - either because of its sample size requirements or its difficulty in just getting a time measurement - can interact with these other factors. Thus we have completed the circuit and probably best summarize this qualitative description of the designing of a work control system with a diagram:



To this point, the description has been purely qualitative. Let us now consider some statistical and mathematical techniques that can aid in the analytical process. We'll start with the more familiar area of designing the measuring process and then move on to selecting the mathematical model. In this part of the discussion, the purpose is to indicate where some well-known procedures can be brought to bear in some - not all - phases of Work Measurement. These techniques will be described by name and by characteristic formulas easily recognizable to those familiar with statistics. References containing complete information on why these procedures work are noted and suggested for those unfamiliar with statistics who want to know what goes on "under the hood."

Designing the Measuring Process

Measurement is the business of getting a number which reflects what is going on. The measuring process has two problems facing it: (1) Accuracy - its representativeness; (2) Precision - its ability to reproduce measurement values. Accuracy depends on the typicalness of the sample that is selected. Random sampling guided by random number tables is an accepted manner of getting a random sample. This subject is well covered in statistical texts (1) and, with regard to Work Measurement in references (2), (3), and (4).

Precision depends on sample size. Here, specific statistical guides are available and are covered in the indicated references. Some of these techniques in the Work Measurement context are summarized below considering first of all variables - or stop watch - work measurement.

Stop Watch Work Measurement - Simple Repetitive Cycle

The approximate 95% confidence interval formula in variables measurement is $2s_{\bar{\bar{X}}} = \frac{2s}{\sqrt{N}}$. This can be simplified by using $\frac{\bar{R}}{d_2}$ in place of s, standardizing on sub groups of 4, (n = 4, so $d_2 = 2.059$) and using a percent precision value $E = \frac{2s_{\bar{\bar{X}}}}{\bar{\bar{X}}}$ The simplified formula:

$$E = \frac{97\,\bar{R}}{\bar{\bar{X}}\sqrt{N}}$$

where E = percent precision
$\bar{R}$ = average range from groups of 4
$\bar{\bar{X}}$ = average cycle time
N = sample size

The engineer can specify E, use a preliminary study, of, say, 16 observations to get $\bar{R}$ and $\bar{\bar{X}}$, and solve for N. Or he can, knowing N, $\bar{\bar{X}}$ and $\bar{R}$ from a study, solve for E. A nomograph for this is displayed in reference (2).

Stop Watch Work Measurement - Complex Cycle

Not all repetitive operations are so simple, however. Consider one with this mathematical model:

Average total time per piece =

Average time per piece + $\frac{\text{average tray handling time}}{\text{average pieces per tray}}$

$$\bar{\bar{X}}_T = \bar{\bar{X}}_A + \frac{\bar{\bar{X}}_B}{\bar{\bar{F}}_B}$$ where X_A, X_B and F_B are all varying

As before, the interest is in the percent precision $E = \frac{2s_{\bar{\bar{X}}_T}}{\bar{\bar{X}}_T}\,100.$

A simplified formula for obtaining E has been developed (2) considering that $s^2_{\bar{\bar{X}}_T} = s^2_{\bar{\bar{X}}_A} + s^2_{\frac{\bar{\bar{X}}_B}{\bar{\bar{F}}_B}}$ and using the customary expression for the variance of a quotient (5) so that the second term becomes

$$\frac{s^2_{\bar{\bar{X}}}\,\bar{\bar{F}}^2_B + s^2_{\bar{\bar{F}}}\,\bar{\bar{X}}^2_A}{\bar{\bar{F}}^4_B}$$. The simplified formula is $E = \frac{97\,R^*}{\bar{\bar{X}}_T\sqrt{N_A}}$ with E and $\bar{\bar{X}}_T$ as defined above and N_A the observations on the main cycle times X_A. $R^* = \bar{R}_A\,(1 + .15u)$ where $\bar{R}_A$ = average range from groups of 4 X_A's and u = the number of auxiliary elements such as B in this case (u = 1 of course, in this case). As with the simple cycle, the engineer can specify a desired E, get $\bar{\bar{X}}_T$ and R* from a short preliminary study, and solve for N_A. Or, having N_A, $\bar{\bar{X}}_T$, and R* from a study, solve for E. The nomograph of reference (2) also works here.

Work Sampling (attributes measurement) - Simple Case of % Busy

The approximate 95% confidence interval formula in attributes measurement is $2s_{\bar{p}} = \frac{2\sqrt{\bar{P}(1-\bar{P})}}{\sqrt{N}}$. In work sampling, $\bar{P}$ becomes the percent busy, $2s_{\bar{p}}$ the precision expressed as an absolute percent plus-or-minus interval, and N the number of busy - not busy observations. The engineer can obtain $\bar{P}$ from a short preliminary study - say 80%, specify $2s_{\bar{p}}$ - say 4% or from 76% to 84%, and solve for N which would be 400 in this case. Or, having $\bar{P}$ and N from a study, he can solve for $2s_{\bar{p}}$. The nomograph of reference (3) handles this calculation.

Work Sampling (attributes measurement) - Complex Case

Quite often the engineer is interested in breaking down percent busy into 2 or more classifications. The mathematical model becomes $\bar{P}_t = \bar{P}_1 + \bar{P}_2 + \ldots$ etc. with the $\bar{P}$ values varying. A simplified formula has been developed based on reference (6) solving for

$$E = \frac{200}{(\bar{P}_1 + \bar{P}_2)\sqrt{N}} \sqrt{\frac{\bar{P}_1'^2(1-\bar{P}_1')}{\bar{P}_1} + \frac{\bar{P}_2'^2(1-\bar{P}_2')}{\bar{P}_2} - 2\bar{P}_1'\bar{P}_2'}$$

In a multi-element study with many $\bar{P}$ values $\bar{P}_1$ = the $\bar{P}$ value with the highest relative upward shift (maximum $\frac{\bar{P}_i'}{\bar{P}_i}$) and $\bar{P}_2 = \bar{P}_t - \bar{P}_1$. A $\bar{P}'$ value is an estimate of the maximum probable $\bar{P}$ value. So, given current $\bar{P}$ values from a preliminary study, estimating $\bar{P}'$ values from the study or history, and specifying E - the desired precision on some future possible mix of proportions, the engineer can solve for N - the sample size to take now. Or, given N, $\bar{P}$ values, and $\bar{P}'$ values, he can solve for E.

Selecting the Mathematical Model

Internal Variation: Recalling the earlier example with the model $T = N\bar{\bar{X}}$, the engineer must be satisfied that this "one unit of measure" model is all right. A first check can treat the average decision period production as a sample and see if the 95% confidence interval on that sample is within some specified tolerance. He essentially solves for a value we'll call $E_r = \frac{2s_{\bar{\bar{X}}}}{\bar{\bar{X}}} 100$ where $2s_{\bar{\bar{X}}} = \frac{2s_X}{\sqrt{N_p}}$ and N_p = decision period production. If $E_r > 5\%$, say, then one unit of measure is insufficient. The expanded model $T = \bar{\bar{X}}_1 N_1 + \bar{\bar{X}}_2 N_2$ could then be tested and so forth. The simplified formula for this using groups of 4, $\bar{R}$, and d_2 is $E = \frac{97\,R^*}{\bar{\bar{X}}_T\sqrt{N_p}}$ and it can be handled on the nomograph of reference (2).

Between Decision Period Variation: Besides short term sampling variation in the operation, there may be inherent decision-period-to-decision-period variation. Determining whether or not this variation (call it $2s_d$) is greater than some specified tolerance (say 5%) can be a tricky

problem. One way is to observe several decision periods completely and get a sample between-decision-period standard deviation. This can be expensive, however, and using the notion of variance components would appear much better. In observing samples from several decision periods (at least 2) the following relationships would pervail:

Total variance = sampling variance + decision period variance

$$s_T^2 = s_s^2 + s_d^2 \qquad \text{and} \qquad s_d^2 = s_T^2 - s_s^2$$

Total variance, s_T^2, is observed; sampling variance, s_s^2, can be calculated by pooling decision period variances; and one can compute s_d^2 - the variance associated with the decision period variation that would prevail if there was no sampling. Components of variance are covered quite extensively in reference (1).

The interest in this case is in $E_v = \frac{2s_d}{\bar{\bar{X}}_T}\ 100$. Getting s_d^2 and then s_d from a preliminary study along with $\bar{\bar{X}}_T$ the engineer can match the computed E_v against the tolerance. If it's greater, then more units of measure are needed and an expanded model can be tested. There is a possibility, of course, that sample size is insufficient for any decision - $s_s^2 > s_T^2$ - and further study with a larger sample is then required.

Multiple Regression - Variables: We have talked to date about examining progressively expanded models until the residual variation is within tolerance. The tacit assumption has been that times for each newly considered auxiliary element of the operation were obtainable by direct observation. This is not always possible such as in the case of packaging operations. The time to package one type of item among many types being packaged is often thoroughly mixed in with and dependent on the times of packaging these other types. All that is known about the operation is that for each package some total time has been put into packaging and a variety of countable types of items packaged. The question is: How many of the item types should be counted for an adequate control system? The question, of course, only comes after having decided by previous tests that just counting packages is insufficient.

This is a problem in multiple linear regression. Consider this typical example wherein from a short preliminary study times for packaging one order (X_T) have been obtained along with the associated counts of the three types of items being packaged (N_1, N_2, N_3).

Time to Package One Order X_T	Number of Items of Type 1 N_1	Number of Items of Type 2 N_2	Number of Items of Type 3 N_3
.92	11	4	5
.67	1	13	1
.47	4	5	0
'	'	'	'
'	'	'	'
'	'	'	'

The results of a multiple regression analysis (see reference (7)) on the complete data from studying 16 orders showed the following:

$$X_T = \bar{\bar{X}}_0 + \bar{\bar{X}}_1 N_1 + \bar{\bar{X}}_2 N_2 + \bar{\bar{X}}_3 N_3$$

i	$\bar{\bar{X}}_i$	Decision Period (Daily) Variation (E_r)
	.541	6.0%
0	.318	-
1	.852	5.9
2	.813	1.2
3	.013	.9

$$X_T = .318 + .852 N_1 + .813 N_2 + .013 N_3$$

Note the recording of the residual variation (S_{y_i}) opposite the associated regression coefficient ($\bar{\bar{X}}_i$). This is phrased in terms of chance variability of a decision period (daily) sample N_p: $E_r = \dfrac{2 S_{y_i}}{\bar{\bar{X}}_T \sqrt{N_p}}$

If ±5% were satisfactory for E_r then item type 3 could be dropped. Subjecting the final study to a two type regression analysis would probably generate a model satisfactory for the control system. Incidentally, sample size can be approximated in such cases by using the previous confidence interval formula $E = \dfrac{2 S_{y_i}}{\bar{\bar{X}}_T \sqrt{N}}$ where S_{y_i} is the cut-off S_{y_i} value (S_{y_2} in this case). Also note that E_v - between decision period variation - can be examined by analyzing the individual regression residual values.

Multiple Regression - Work Sampling: The same approach can be used with work sampling results from a non-repetitive operation study. A work sampling percent busy can be easily converted to a time estimate. Consider these data from a study of the electrical trades:

Job No.	Percent Busy		Total Job Time (from job tickets)		Direct Labor Time
1	73	x	52 hours	=	38 hours
2	65	x	171 "	=	111 "
'	'		'		'
'	'		'		'
'	'		'		'

The time inputs for each job studied and the associated counts for each job of the conduit sections, footage of small size conduit installed, and footage of large size conduit installed - these were the ingredients for the multiple regression analysis:

Job No.	Work Sampling Hours X_T	Conduit Sections N_1	Small Size Conduit Feet N_2	Large Size Conduit Feet N_3
1	7.3	17	13	72
2	76.7	180	804	241
3	9.0	29	0	62
4	26.7	73	399	0
'	'	'	'	'
'	'	'	'	'

$$X_T = .25 + .12\ N_1 + .04\ N_2 + .08\ N_3$$

$2\ S_{y_3}$ = .3 hours or ±4% on the average job

Fewer or additional units of measure could have been considered but this residual variation was felt to be correct.

Queueing Models: Service type labor is often analyzed in Work Measurement. This is labor that is on a demand basis - labor of an indirect or non-repetitive type whose services are constantly on call but only intermittently, randomly used. Machine-breakdown repairmen are in this category and are increasingly important as automation moves into the picture. Oddly enough, this control problem has become more prominent as the use of work sampling has increased. For work sampling has effectively disclosed, time and again, the fact that a service crew is inactive part of the time. But this information is not enough. Arbitrary crew reduction merely creates machine delays on those occasions when several machines demand service at once, possibly as costly as idle labor.

The idle labor can be considered as labor "queueing up" or waiting in line to service machines. Machine delays can also be considered as machines "queueing up" to be serviced. This queueing problem can be resolved by the use of queueing theory - a segment of probability theory - which permits the engineer to predict the effect of various crew sizes on the idle labor and machine wait time characteristics.

Here is a simple example. At our Kodak Park Works, the Kodacolor Printing Room, at a particular production level, had a crew of 3 handlers per shift supplying and inspecting magazines of photographic paper for 28 printers. Since these magazines had paper of varying length and the printers had varying speeds, the demand pattern for supplying magazines was quite unsystematic. Nobody knew just when a printer, or, indeed, several might be in need of a magazine change. The inspection of magazines was routine, however; they could be inspected at leisure and stored until needed. So the problem was one of controlling non-repetitive labor, a portion of whose services were randomly demanded.

Work sampling results showed the three men were busy 47% of the time (1).

The obvious move to reduce the crew to 2 men and experience about 71% activity (47% x 3/2) was not automatically the right move owing to the demand nature of their work. Queueing theory had to be employed to evaluate the effect of such a possible move.

The engineer procured the following basic information about the operation:

1. $m = 28$ = number of printers under consideration
2. $r = 2$ = number of servicemen under consideration
3. $a = \frac{\ell}{u} = .0246$

Where $\ell = 3.1$ minutes = the average time for a handler to supply a magazine to a printer.

$u = 125.8$ minutes the average run time of a printer

The average magazine change time (ℓ) was obtained from regular stop watch studies. The average printer run time (u) was obtained from printer logs where the paper footage for each run was recorded. Knowing printer speed, time for each run and the average run time were simple to compute. With this basic information, the engineer used a table based on this mathematical model (see reference 8):

$$\frac{P_n}{P_o} = \frac{m!\, a^n}{n!\,(m-n)!} \qquad r \geq n \geq o$$

$$= \frac{m!\, a^n}{r!\, r^{n-r}\,(m-n)!} \qquad r \leq n$$

$$P_o = \frac{1}{\sum_{n=o}^{m} \frac{P_n}{P_o}}$$

$$P_n = P_o \left(\frac{P_n}{P_o}\right)$$

$$i = \frac{\sum_{n=r+1}^{m} (n-r)\, P_n}{m} \qquad r \leq n$$

$$f = \frac{\sum_{n=0}^{r} (r-n)\, P_n}{r} \qquad n \leq r$$

where P_n = the probability of n machines being down together (requiring service)

($P_o = P_n$ for $n = 0$)

i = the average machine wait time expressed as a decimal fraction
f = the average serviceman wait time expressed as a decimal fraction

He determined the values as follows:

i = .0033 or about .3%
the average fraction wait time per printer

f = .6353 or about 63.5%
the average fraction wait time per handler

Thus he could say that if the crew were reduced to 2 men, the delay introduced by such a reduction (delay due to interference) would be about .3%. The handler idle time, of course, would be much less than 63.5% in actuality owing to the inspection portion of the job; the work sampling 2-men crew estimate of 71% busy, or 29% idle, would still hold. Queueing theory only considered the random demand portion of the operation. A final check on the aplicability of the model had to be made. The computed i and f values for the observed (3 man crew) situation were compared to the actual sampling results on i and f:

	Observed	Calculated by the Model
i	0	.03%
f	73% ± 5%	75.6%

These observed data fitted the model for all practical purposes meaning the 2 man crew prediction was quite likely valid.

There are other machine interference models (reference 9) handling other types of situations. Note that here we can achieve improvement, perhaps, in the operation as well as gain control. Production level shifts (more machines) can be considered in this model and, for a desired i and f balance, the optimum crew size assigned. The "unit of measure" here is essentially the number of machines and crew size is the dependent variable. The function is not a simple linear one but can be tabulated for scheduling use. Many "indirect" type operations are susceptible to this type of control where the unit of control is an individual - not some portion of his or her time as in simple repetitive cycle.

Linear Programming and Handling: Considerable attention has been paid to the "indirect" operation of handling. Work sampling has made it possible to measure this activity but activity measures do not always resolve the problems of either control or improvement. The linear programming model is effective in specifying the optimum deployment of handlers to meet specified service requirements. Consider this handling problem with the following demand matrix of loads per hour:

		From		
	Area →	1	2	3
	↓			
	1	0	12	4
To	2	6	0	2
	3	5	1	0

These areas are equidistant. If a handler can move 6 loads per hour then the ideal number of handlers in all is $\frac{6 + 5 + 1 + 12 + 4 + 2}{6} = \frac{30}{6} = 5$.

To set up a good method of handling (not necessarily the "one best method" but a good one) by linear programming to achieve one hour service some possible routes can be listed - say 10 to start with. And associated with each route is an unknown number of handlers:

Route		Number of Handlers
I	1, 2, 1, 2, 1, 2, 1	a
II	1, 3, 2, 1, 2, 3, 1	b
III	1, 3, 1, 3, 1, 3, 1	c
	etc.	

Now, for each element in the demand matrix an inqeuality can be written. For instance, element 1, 2 requires 6 loads per hour be moved. Route I can move 3 loads for each handler on that route and Route II can move one load for each handler it has. So, $3a + b \geq 6$ if these 3 routes are the only ones considered. Other routes among the initial 10 might have a 1,2 segment so the general inequalities for this and all elements become:

Element

1,2	3a	+	b	+	.	.	.	.	.	$\geq$	6
1,3			b	+	3c	+	.	.	.	$\geq$	5
3,1			b	+	3c	+	.	.	.	$\geq$	4
					'						
					'						
					'						

and we would want to minimize the total number of handlers $Z = a + b + c + \ldots$ These inequalities along with the function to be minimized constitute a linear programming problem. Solution techniques, some underlying ideas in linear programming and a report on a somewhat similar application with railway freight cars are contained in the references under (10).

A few comments on this type of application - the solution to the above may not be the best since all possible routes have not been considered. However, it's very easy to write down any number of additional routes and assess their improvement, if any. Also - high variation in demand may upset service. Each channel with a given schedule can be examined with queueing theory and the channel especially loaded if the waiting line is too long. Note that here, again, both improvement and control are achieved. As demand increases or changes the deployment of handlers must change. The relationships can be pre-computed and tabulated - the simple linear relationship of production x average time no longer holds.

Summary

There are four basic factors affecting the design of a work control system: The decision tolerance - how close you want to be in predicting or stating labor content; the decision period - how frequently you want to make a prediction or statement; the mathematical model - symbolic portrayal of the elements in the operation; and the measuring process - what the engineer has to do to get the numbers for the mathematical model. These factors all interact and the industrial engineer seeks a balance among them every time he undertakes a Work Measurement study. This balancing is usually intuitive and usually expensive in that poor decisions require later analysis to repair the situation. Statistical and mathematical techniques do exist to help the industrial engineer handle these problems with considerably greater effectiveness. As demonstrated in Quality Control and in Operations Research, the payoffs can be huge on both cost cutting for the plant's operations and reduced engineering time.

A vast gap exists, however, between the knowledge available and the knowledge used. Quality Control people are ideally situated in industry to bring these ideas to others' attention and to demonstrate, by their own efforts in their own field, that there is a vast potential. In proportion as we all team up and tap the immense theoretical power that is just out of reach now, we can move toward our mutual goal of bringing more science to management.

References

1. Dixon, W. J., and Massey, F. J., Jr., Introduction to Statistical Analysis, McGraw-Hill, New York, 1951.

2. Allderige, J. M., "Statistical Procedures in Stop Watch Work Measurement", Journal of Industrial Engineering, August-September, 1956.

3. Allderige, J. M., "Work Sampling Without Formulas", Factory Management and Maintenance, March, 1954.

4. Barnes, R. M., Work Sampling, U.C.L.A., Los Angeles, 1956.

5. Arkin, H., and Colton, R. R., Statistical Methods, Barnes and Noble, Inc., New York, 1955.

6. Cramér, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, New Jersey, 1946.

7. Croxton, F. E., and Cowden, D. J., Applied General Statistics, Prentice-Hall, Inc., New York, 1953.

8. Allderige, J. M., "Tackling Old Problems With New Tools From Operations Research", Advanced Management, September, 1956.

   Feller, W., Probability Theory and Its Applications, John Wiley and Sons, Inc., New York, 1950.

9. Anson, C. F., "Determining Machine and Operator Utilization and Synchronization Allowances", Time and Motion Study, November, 1956.

   Fetter, R. N., "The Assignment of Operators to Service Automatic Machines", Journal of Industrial Engineering, September-October, 1955.

10. Charnes, A., and Miller, M. H., "A Model For Optimal Programming of Railway Freight Train Movements" paper presented at the annual meeting of the Econometric Society, December 30, 1955. Management Sciences Research Group publication, Purdue University, Lafayette, Indiana.

    Churchman, C. W., Ackoff, R. L., and Arnoff, E. L., Introduction to Operations Research, John Wiley and Sons, Inc., New York, 1956.

QUALITY CONTROL APPLIED TO INDUSTRIAL ENGINEERING

Charles R. Hicks
Purdue University

As I have been assigned one-quarter time in the department of Industrial Engineering during this present year at Purdue, I hope to be able to give some up-to-date examples of the application of statistical methods to industrial engineering problems. I would prefer to generalize the topic somewhat to include many statistical methods and not just those which I usually associate with quality control. I prefer to think of quality control as somewhat restricted to those techniques directly aimed at control of the quality of the product such as control charts, acceptance sampling plans and the like. I realize that many quality control people often consider quality control as covering any and all statistical techniques applied to industrial problems. Our topic might be stated then as "Statistical Methods Applied to Industrial Engineering".

Of all the branches of engineering, probably the industrial engineer has been most receptive to the use of statistical techniques although other branches of engineering are gradually beginning to appreciate the value of statistical and probabilistic thinking. Evidence of this is found in the fact that nearly all accredited Industrial Engineering curricula in our universities now require at least one, and usually two, courses in statistics and quality control for their undergraduate engineers.

Now why is the industrial engineer so concerned with statistics? Let us look at a definition of Industrial Engineering adopted recently (1956) by the American Institute of Industrial Engineers (A.I.I.E.):

"Industrial Engineering is concerned with the design, improvement, and installation of integrated systems of men, materials and equipment. It draws upon specialized knowledge and skill in the mathematical, physical, and social sciences together with principles and methods of engineering analysis and design to specify, predict and evaluate the results to be obtained from such systems".

From this definition, we can see that Industrial Engineering is concerned with the design of systems involving men, materials and equipment and with analysis and evaluation of results obtained from these systems. Surely statistical techniques are necessary here in designing experiments and in analysis of the results. The A.I.I.E. was started in 1948 and is thus a contemporary to A.S.Q.C. It now has 53 chapters and some 4000 senior and associate members and many of its activities and interests overlap those of our society. In the last few years with the advent of this undefinable monster called "Operations Research" the science of Industrial Engineering has one of its greatest opportunities to grow and develop more professional status than it has had in many years. Many of the techniques of operations research are 'naturals' for the industrial engineer but do require a stronger background than he has been used to in mathematics, statistics and probability theory. One of the writers in the area of Industrial Engineering stated that "two needs confront Industrial Engineering education if we are to make the transition to professional practice. These are:

1. The introduction of more powerful quantitative, mathematical, statistical methods to provide more precise tools for planning and evaluation.

2. The introduction of and thorough grounding in the behavioral and social sciences to prepare industrial engineers for the integrated systems design job they must carry out."

Now as to specific applications of statistical tools in Industrial Engineering, I believe a paper by Professor Howard P. Emerson of the University of Tennessee, and present President of A.I.I.E. entitled "A Mathematics Foundation for Industrial Engineering"(1) is the best single source of information on such applications. Professor Emerson spells out in two columns the statistical techniques and the corresponding Industrial Engineering application. Some such applications are Quality Control, which we are all quite familiar with as it applies to control of process variables; Motion and Time Study, which is often given the more sophisticated name of Work Measurement, Merit Rating, Plant Layout and Design, Estimating and the Precision of the Estimate, Methods Improvement, Evaluation of sources of variation, Operations Research, etc. To discuss all of these and many other applications of statistical methods to Industrial Engineering would far exceed our time limitation, so I should like to discuss in some detail three problems from the areas of Work Measurement, Experimental Design and Replacement Theory the latter being often construed as belonging in the Operations Research area.

I. WORK MEASUREMENT - The problem to be presented is taken in somewhat simplified form from a paper by John Allderige in a recent issue of the Journal of Industrial Engineering.(2)

The problem was to determine how large a sample (n) should be taken for a time study to determine the average length of time for a given cycle, how to be assured that the results obtained are consistent, and what precision can be expected in the result or how close may the result be to the true average time. This problem has been attacked by many people in Industrial Engineering but Allderige makes use of many of the ideas we use in control chart analysis which should make his approach of interest to the quality control man.

The method proposed is to take a preliminary sample of 4 groups of 4 readings each of the time to complete the given cycle, these 4 groups of readings being taken at random times throughout the day. Using the 4 groups, 4 averages and 4 ranges are computed and from these the grand average ($\bar{\bar{X}}$) was found (for the problem he gives, equal to 1.00 minutes) and the average range ($\bar{R}$) = 0.40 minutes. With this information and a statement as to the precision desired in the final average time of the cycle (say, to be within 5% of the true average time), he determines the number of additional cycles to be timed. This additional sample size is given by means of a nomograph in the paper but can be worked out for the data given as follows:

Assuming a normal distribution of time study data the statistic:

$$Z = \frac{\bar{\bar{X}} - \bar{\bar{X}}'}{\frac{\sigma'}{\sqrt{n}}}$$ is normally distributed

with mean of zero and standard deviation of one. Here $\bar{\bar{X}}'$ is the true average time we wish to estimate and $\bar{\bar{X}}$ is the observed average time. We want $\bar{\bar{X}}$ to be within 5% of $\bar{\bar{X}}'$ or $\bar{\bar{X}} - \bar{\bar{X}}' = .05\bar{\bar{X}}'$. As $\bar{\bar{X}}'$ is unknown we

use $\bar{\bar{X}}$ as the best estimate of $\bar{X}'$ based on the original data. σ' is the true standard deviation of the process and can be estimated by $\bar{R}/d_2$ as is done in quality control chart analysis. Z is usually set at 1.96 or 2 representing two standard deviations on a normal distribution which will give approximately 95% confidence that our n will be sufficient to estimate the true mean with 5% precision. If the expression above is then solved for n, we have the total number of cycles that should be timed. Or, for our data:

$\bar{\bar{X}} = 1.00$ min. $\qquad \sigma' = \bar{R}/d_2$

$\bar{R} = 0.40$ min. $\qquad = \frac{.40 \text{ min.}}{2.059}$

$\bar{X} - \bar{X}' = .05\bar{\bar{X}}$

$Z = 2$ $\qquad$ as $d_2 = 2.059$ for samples of 4

Hence substituting:

$$2 = \frac{(.05)(1.00)}{\frac{.40}{2.059\sqrt{n}}}$$

Solving this for n, we get n = 61. As we have already taken 16 observations, this means to take 45 more. To check on the consistency of the results, an $\bar{X}$ chart is suggested; which, if in control for the 15 or 16 points (if we continue to take groups of 4), means that the average time is consistent from hour to hour throughout the period of the chart. Thus a check can be made on consistency and the precision required can be set in advance. If, for example, we had wanted 1% precision, we would have used $\bar{X} - \bar{X}' = .01\bar{X}$ and the resulting n would have been 1509 observations or 1493 more than the preliminary sample!

This particular approach is quite simple compared to some others that are used as it uses ranges rather than standard deviations. The only serious question about this approach is the assumption of normality of time study data which is being studied at Georgia Tech. and Purdue at the present time. There is a strong indication that time study data may be positively skewed.

II. EXPERIMENTAL DESIGN - Designing an experiment to study the effects of men, materials and equipment is very important to the industrial engineer. In the problem to be discussed, one of our graduate students at Purdue wished to study the effect of the degree of symmetry of an object on the time required to position and assemble two mating parts. Previous studies suggested that objects could be classified as perfectly symmetrical (circular), non-symmetrical (fitting only one way) and semi-symmetrical (all others). An experiment was designed to determine whether an increase in the degree of symmetry from non-symmetrical (or 1 degree of symmetry) up through 6 degrees on symmetry would tend to decrease the time required to position and assemble two mating parts. The experimenter chose to study profiles having one, two (rectangle), three (equilateral triangle), four (square) and six (hexagon) degrees of symmetry and see how the average time to position and assemble two mating parts with these profiles would change. In planning the experiment it was readily recognized that there were several other variables which would effect the average time to position and assemble a part, the most obvious one being the operator who performed the task. It was decided to select 14 operators at random from graduate industrial engineering students at Purdue and try several runs using each of the 5 different profiles. The profile presented to a given operator first, second,

third, etc. was decided by a random presentation in order to help reduce the effect of learning and fatigue. As weight of the object might be an important variable affecting the time, all profiles were hollowed out such that all had approximately the same weight. It was also agreed to have each operator try several runs using each of the 5 profiles and then only the middle 10 readings were used in the analysis. These and other considerations led to the formulation of a mathematical model to describe the experiment. One of the more modern trends in the operations research area is to try and formulate a problem as a mathematical model. Our model here is given by:

$$Y_{ijk} = \mu + P_i + O_j + (PO)_{ij} + \epsilon_{k(ij)}$$

where: Y_{ijk} is the observed time for the k^{th} run of operator j using profile i.

μ is the common mean in all data.

P_i is the profile effect where i = 1, 2, 3, 4, 5 for the 5 profiles.

O_j is the operator effect where j = 1, 2, . . . 14 for the 14 operators.

$(PO)_{ij}$ is a term representing a possible interaction between profiles and operators.

$\epsilon_{k(ij)}$ is the random error as seen in the k runs for each operator, profile combination. Here k = 1, 2, . . . 10 runs.

Such a model implies a factorial design with 5 x 14 x 10 = 700 observations and the results to be analyzed by the analysis of variance technique. Another feature of this design is that it is a mixed model in that the profiles are fixed and the operators are random.

As the analysis of 700 observations seemed quite laborious, several short cut statistical techniques were used which should be of interest to quality control people. For each of the 70 cells (5 profiles by 14 operators) the average and range of the 10 observed times were computed. The ranges were then plotted on a range chart and this was examined for homogeniety of variance within the 70 cells. The more exact test would be a Bartlett test which would be much more work. Also from the average of these 70 ranges which were in good control an estimate was made of the standard deviation of the errors using $\bar{R}/d_2$ with d_2 based on samples of 10. Then the analysis of variance was run using only the averages in each cell as the variable. This, of course, gave no exact test for interaction but this could be checked from the estimate of error made by using $\bar{R}/d_2$. Whether the interaction is significant or not will not affect the test on profiles as this is a mixed model. It was decided to test at the 1% level of significance and the following table was compiled from the data:

ANALYSIS OF VARIANCE TABLE ON PROFILES

TABLE 1

Source of Variation	d.f.	Sum of Squares	Mean Squares	Expected Mean Square
Profiles	4	6,224.51	1,556.13	~~$\sigma_e^2+\sigma_{PO}^2+14\sigma_P^2$~~
Due to linear reg.	1	4,350.19	4,350.19	
Departure from lin. reg.	3	1,874.32	624.77	$\sigma_e^2+5\sigma_O^2$
Operators	13	18,124.00	1,394.15	
O x P and error	52	12,229.74	235.19	$\sigma_e^2+\sigma_{PO}^2$
Total	69	36,578.25		

Looking only at the profile, operators, and P x O and error line, the expected mean square column shows that the proper test for profile effect on the average time is to compare the profile mean square with the O x P and error line mean square. Thus the F ratio here is:

$$F(4,52 \text{ d.f.}) = \frac{1556.13}{235.19} = 6.62$$

which shows a significant profile effect at the 1% level of significance. This gave pretty good evidence that the degree of symmetry as revealed by these 5 profiles has a very real effect on the average time necessary to position and assemble mating parts.

As this test indicated that the degree of symmetry did affect the time, it was then wondered what the relationship might be. Plotting the average times for each degree of symmetry led to the hypothesis that there might be a linear relationship between degrees of symmetry and time with time decreasing as the degree of symmetry increases. From the data the sum of squares due to linear regression was also computed and is given in the table above. The difference between this and the sum of squares among profiles is due to the departure from linear regression and this departure was tested by comparing the mean square of 624.77 with the 235.19 or:

$$F(3,52 \text{ d.f.}) = \frac{624.77}{235.19} = 2.66.$$

This is not a significant departure from linear regression at the 1% significance level so that a linear relationship might be reasonably assumed and the equation of such a line could be written. We also checked the value of eta squared = 0.34 and r^2 = .24 which indicates that about 24% of the variance in average times might be accounted for by the straight line and 34% might be accounted for by a line through all 5 average points. (For the benefit of the astute statistician, we subtracted the operator sum of squares from the total before calculating these statistics).

More research is in progress on this problem but this example is presented to give some indication of the kinds of statistical and quality control techniques that are useful to an industrial engineer.

III. REPLACEMENT THEORY - A problem reported by some workers at Case Institute of Technology has to do with when to replace items in the plant. Is it best to replace a machine, say, when it breaks down or is

it better to replace several machines at regular intervals regardless of their condition? If the latter is best, what is the proper replacement period in order to make the overall cost a minimum?

The problem cited is a somewhat simpler problem on the replacement of a system of 10 light bulbs. It is assumed that it costs $1.00 to replace a bulb when it fails and this must always be done. It costs only 50 cents per bulb when replacing all 10 bulbs at one time. The problem then is to decide on a proper replacement period; say 10 days, 20 days, 30 days or 40 days to replace all 10 in order to give a minimum annual cost realizing, of course, that any single bulb will have to be replaced whenever it fails.

A mathematical model giving the total cost is expressed as follows:

$$C = \frac{T}{t_r} [N \cdot P(t_r) C_1 + N \cdot C_2]$$

where C_1 = cost of replacing an individual bulb = $ 1.00
C_2 = cost of replacing an individual bulb if replaced as one of a group of 10 = 0.50
N = number of units in the system = 10
t_r = time between group replacements = 10, 20, 30, 40 days
C = total cost of this replacement policy over time T.
$P(t_r)$ = probability that a bulb will fail before time t_r.

Every variable in this equation is known except the probability that a bulb will fail before time t_r. In order to estimate this probability for time periods (t_r's) of 10, 20, 30 and 40 days, a Monte Carlo technique is suggested. This simply means that we play a game with some 'typical' data to empirically estimate the value of this probability for each time period t_r. This can be done without any expensive experimentation if we have some past data on light bulb failure.

Assuming that bulb life is approximately normally distributed with a mean bulb life of 30 days and a standard deviation of 10 days, a table of normally distributed random numbers can be used to assign a random life to each bulb. If, for example, the normal random number drawn is 0.5, this means that the bulb would be assigned a life 0.5 standard deviations above the mean or 35 days. Such a bulb would survive a 10, 20, or 30 day replacement but a new bulb would have to be put in before the 40 day period. In this latter case, a new random number would be drawn to see if the new bulb would survive the remaining 5 days until the 40 day replacement. Sometimes several bulbs may have to be used in a given socket before the 40 day period occurs.

Such a 'game' was played for the 10 bulb sockets with some of the results shown below:

BULB FAILURES FOR A GROUP OF 10 BULBS

TABLE 2

t_r Bulb Socket	1	2	3	4 . . .	. .	10	Total Failures
10 days	35	31	55	27			0
20 "	35	31	55	27			0
30 "	35	31	55	27,55			4
40 "	35,66	31,36,62	55	27,55			9

The numbers entered in the body of the table are the number of days life for each bulb chosen at random. For socket 1, for example, the first bulb with 35 days life would survive the 10, 20, and 30 day replacement but not the 40 day replacement. Hence, another bulb was inserted (drawn from our normal random table) and its life was 31 days which when added to the 35 gave the 66 days shown in the table which means that the 2 bulbs would last until the 40 day replacement. In the case of the second socket, however, the second bulb chosen had a life of only 5 days making 36 in all which means still another replacement before the 40 days replacement period. Adding the number of replacements for each period for the 10 sockets gave the values shown at the right of the table. One trial of the "game" is hardly enough to estimate the probability of failure before time t_r so the game was repeated 10 times giving the following number of failures for each trial:

BULB FAILURES IN 10 TRIALS OF 10 BULB EACH

TABLE 3

Time t_r	1	2	3	4	5	6	7	8	9	10	Total Failures	$P(t_r)$
10	0	0	1	0	0	0	0	0	0	0	1	0.01
20	0	1	3	3	2	2	2	1	0	1	15	0.15
30	4	4	8	5	8	6	5	5	3	3	51	0.51
40	9	10	11	11	9	10	12	8	9	7	96	0.96

If the total failures in 10 trials of 10 bulbs per trial is now divided by 100, the last column gives a pretty good approximation to the probability of failure of a single bulb for time t_r. If these probabilities along with the corresponding t_r values are now substituted in the cost equation, we find:

$$C = \frac{T}{t_r} \left[N \cdot P(t_r)\, C_1 + NC_2 \right]$$

For 10 day replacement: $C = \frac{360}{10} [10(.01)(1.00) + 10(.50)] = \$183.60$

20 day replacement: $C = \frac{360}{20} [10(.15)(1.00) + 10(.50)] = \$117.00$

30 day replacement: $C = \frac{360}{30} [10(.51)(1.00) + 10(.50)] = \$121.20$

40 day replacement: $C = \frac{360}{40} [10(.96)(1.00) + 10(.50)] = \$131.40$

A glance at these annual cost figures reveals that a 20 day replacement period would give the least cost to the company. It is interesting to note that this result is obtained by playing a game with some numbers drawn at random from a table of normally distributed random numbers and no expensive experimentation was necessary. Some mathematician might wonder why we didn't simply differentiate the cost equation and find the value of t_r which gives a minimum cost. Because of the probability function involved, this procedure would require taking the limits of an infinite product of normal integrals which is quite prohibitive from a practical point of view. The results above might be improved still further by trying replacement periods around 20 days such as 18, 19 and 21, 22 etc. and use a Monte Carlo scheme of game playing to narrow the results down to the best number of days for the 10 bulb replacement. This very practical problem for an industrial engineer has been included to show how a simple application of normal random numbers can be used to help solve the problem.

Many other examples might be mentioned illustrating the vital role of statistics in Industrial Engineering but these three should give some indication of the kinds of applications that have been made. At the present time, Purdue industrial engineers are working on such problems as a mathematical approach to plant layout seeking the optimal layout, a statistical analysis of the factors affecting plastic tools, checking the normality assumption of time study data, work on a statistical determination of economic lot size etc.

Most any industrial engineer who is well informed in his field will readily agree that statistical methods provide one of the best tools for solving his problems.

REFERENCES

1. Emerson, Howard P., "A Mathematical Foundation for Industrial Engineering", Jour. Ind. Eng., IV No. 1, February 1953.

2. Allderige, John M., "Statistical Procedures in Stop Watch Work Measurement", Jour. Ind. Eng. VII No. 4, July-August 1956.

WHAT IS AN INSPECTOR

Daniel G. Meckley, III
York Division of Borg-Warner Corporation

History

To talk about the modern concept of an inspector, we must examine the historic and prehistoric beginnings of the breed. The development of the inspector has many elements of evolution and many elements of a modern day revolution. Inspectors are not new; the Bible records instances of overseers and inspectors in terms of rather evil executions of kings' wishes.

Our first concept of inspection was the detection of defective work, largely based on the premise that the operators were trying to slip something over on the company. This quickly led to the operators banding together in an attempt to frustrate the inspector and show that they were smarter by slipping in sub-standard work which he could not detect. Subtly, the operators' entire attitude changed, and in many cases they lost their desire to do the job right.

The real tragedy occurred when the operator mentally shifted the burden of acceptable work from his shoulders to the inspector. Even with the best inspectors, the relationship between them and the foremen and operators was one of an armed truce; the inspector was a friendly spy.

Obviously, this is an exaggerated and over-simplified picture of the pre-World War II inspection setup. Unfortunately, it is truer than we might care to admit. This general situation inevitably led to 100 percent inspection and severe final inspection as attempts to stem the tide of defective work.

Fortunately, the impact of the manpower shortage, and the gradual infiltration of statistics from the few pioneering companies in the field, forced changes to this concept.

I do not believe that any one individual sat down and said that our philosophy of quality control was archaic, but the minds and attitudes of many people, coupled with the facts from the statistical usage, gradually evolved a new concept and philosophy of inspection. This was part of the evolution and development of today's inspector.

The revolutionary changes have come during World War II and since. We now have a different picture of the inspector. Day to day we have had to rapidly shift our thinking to keep up with the technological changes of the last few years.

Whether human or mechanical, the old cliche of "it's here to stay" is applicable to inspection. We can only speculate what its form will be in the future. We want to look at what it is today and ponder over this part of our industrial complex.

To talk about the specific job of an inspector, we will start with a definition of our basic quality task.

The Task

The design and production of complex, high volume, multiple model, appliance type products with yearly changes in style and model places a severe quality responsibility on the operating management.

Figuring strongly in the successful execution of this type business is the ability to design and build "trouble free" products. Products that are immediately "free" from design and manufacturing mistakes which cause production stops and/or field failures. The vechicle for accomplishing this is a well planned, well operated Quality Control system from laboratory through production.

The customers may buy for price and appearance, but consciously or unconsciously they expect their purchases to run trouble free and to have a guarantee in case they do not. Their friends are quickly told about the brands that are "no good".

In the case of the air conditioning business, the York Division of Borg-Warner and other major manufacturers have five year protection plans with their refrigerant circuits; this requires design and conformance quality good enough to run five years without failure. This industry is one that is new enough to have major technical changes coming one on the heels of the other; each one requiring major decisions and risk estimates that must figure into the five year warranty.

A $1.00 mistake on 10% of 100,000 units is a $10,000 corporate liability. A $10.00 mistake on 10% of 100,000 is a $100,000 corporate liability.

Certainly, costs are developed with reserves for guarantee accounts, but increases in these accounts have to be reflected in increased selling price or reduced margins; these are bad dreams to the salesmen and the treasurers.

How "Good" Must The Product Be

I spoke about the product being "free" from defects.

The freedom is relative, and the degree necessary to meet competition successfully is established by the customer, the manufacturer, and the competitor. Each manufacturer designs and builds into his product, and perhaps advertises, a certain quality. Initially the customer doesn't know what a BTU is, but he comes to expect a BTU quality, and gradually learns some of its meaning and other technical measures of quality. He begins to ask for it. The manufacturers then race to get the most BTU's for the least cost.

Eventually a commercial level of BTU acceptability appears as more or less standard. The decision each competiting company finally makes is how much better than barely acceptable does it want to be.

The establishment of a given company's goals in relation to commercial acceptance, service and performance, is the first high level decision necessary before a clean cut practical plan of laboratory to production quality control can be developed.

The control of quality of the type products in question is like hunting; the best opportunities jump up and are quickly gone. If you don't know where your target is located, you can fire box after box of shells without hitting the quality in question. You mised the opportunity because after the first shots the quality becomes more difficult to locate; you have lost the easy killing shot. You shoot some game, but it's more luck then skill.

We all speak best about our own experience so I will take the liberty of telling you how the York Division of Borg-Warner approached the control of quality from the laboratory through production.

What Is The Quality Objective

First, a very general but meaningful definition of the quality goal wanted by the York Division of Borg-Warner was established. We selected: "The best competitive quality with a minimum cost for achieving this position."

This obviously requires further amplification. Take the first part, "The best competitive quality." How is it defined? How is it measured? Who measures it?

How Is The Customers' Satisfaction Measured

We define it as the registered service complaints. The fact that these will contain both problems of design quality and conformance quality, is of no interest to the customer. He sees only defective product; so we measure our performance in his terms regardless of who is responsible.

The measure of quality is percent defective of the installed products. Eight to ten major performance characteristics are measured on each product. These are selected from records and opinions as the performance quality items directly recognized by the customer. The characteristics are common to all air conditioning and ice making equipment so we can chart by products and characteristics.

The data flows daily from all authorized repair men through the Service Department. Here it is collated and sent to Quality Control.

This data presents the state of the unit; but is it good or is it bad?

Quality Control plots the field performance against the target for each characteristic. The target for a given month is calculated to consider the years the product has been installed, and is calculated by assuming that the occurrences of defects are linear with relation to time of installation. The monthly target calculations are based on the five year target.

How Are Quality Targets Derived

The five year targets are developed yearly by Sales, Service, Engineering, and Quality Control. They are approved by the division and corporate management committees. These committees consist of the

division managers and corporate officers. Submission to the management groups gives them a chance to exercise their judgment and experience in the final target establishment.

How do we factually know that these targets will be better than our competitors' performance? We don't, but the persons selecting the targets are in the best position to hear and see about where our competitors are performing; their composite judgment has served our purpose well.

How does the group selecting the targets know what the achievement of these will cost in design and manufacturing? They don't, but since the targets are developed as a must to meet competition in the areas the customer considers important, cost can be held as a second consideration. The cost of major changes required to meet a given target are weighed in terms of tangible lessening of customer complaints, but the decisions to make a major change are not totally accepted or rejected for that reason alone, as frequently customer dissatisfaction makes changes mandatory. Naturally, when large sums are involved, considerable weighing of the tangible and intangible benefits occurs before decisions are rendered.

The effort to achieve the target with a "minimum cost of quality" is the second part of the definition, and also requires explanation.

What Are The Costs of Quality

The costs of quality as we recognize them are:

1. The cost of errors (both salvage and scrap)
2. The guarantee cost
3. The Quality Control Department operating cost
4. The Service Department operating cost
5. The cost of defective purchased material
6. The cost of inspection and test facilities
7. The hospitalization cost (the salvage time allowed in standard costs)
8. Idle time and production losses due to quality problems
9. Clerical cost of processing error replacement orders and salvage orders
10. Cost of direct labor testing
11. Space for inspection, testing and storage of reject material
12. Shipping losses

You may argue or add or subtract in your own case, but we say that the sum of these costs should continuously decrease.

Therefore our approach is simply to reach the performance target and spend less, and less to do it.

Again the intangible or customer satisfaction sometimes upsets the cost of quality. You may be forced to spend more for inspection, a control system, or tooling then the actual saving in guarantee account dollars would indicate is permissible.

We have found these cases to be the exceptions. Most times we

are able to have our cake of desired field quality, and still eat our cake of lower cost of quality.

You may be starting to speculate on the reasons for such elaborate emphasis on the quality goals and their measurement as related to laboratory to production control.

What Is The Quality Control Department's Job

With this general background of the corporate quality objective, we will jump down to the task of the Quality Control Department. We consider that its function should be that of a navigator. The Quality Control Department should act as the agency which provides the information where conformance quality has been, where it is now, where it is trending, and how it can get back on course most quickly.

Its responsibility is to focus the bright light of attention on the quality of the product so that the design and manufacturing time on new products is spent for basic corrections in the place where there is the greatest gain in customer satisfaction and the greatest reduction in the cost of quality.

Inspection Defined

We define inspection as the determination of the acceptability of an item using the established standards as the criterion of acceptance.

The place of the inspection organization in the industrial hierarchy can be argued many ways. We will not concern ourselves with this aspect, but try now to focus our attention on the task of the inspector and his problems.

The Inspectors' Dilemma

Early in this talk I spoke about the cost derived from defective workmanship and product design quality. Consider now the mistakes which can occur from the defective judgment of an inspector.

In the most modern plant there are many facets of the quality which can only be decided by the judgment of individuals. On a fast moving assembly line there is no practical instrument to say that the cabinet on this television set is acceptable or unacceptable from the standpoint of its painted appearance. If, in the inspector's judgment it is defective, the wheels of a vast assembly line will stop. It is seldom that we realize how much responsibility we place in the hands of a relatively untutored $1.80 an hour semi-skilled workman. He can do thousands of dollars of damage in several minutes.

The inspector is both judge and jury as far as the foremen and workmen are concerned. Every human being has certain personality traits and technical deficiencies which at best we can hope to minimize. These reflect themselves in his decisions as judge and jury.

We expect the inspector to represent our corporate viewpoint; yet in most cases he is part of the bargaining unit.

We expect him to be flexible, but we demand that he does not compromise the quality of the product. In a sense we demand both flexibility and inflexibility.

We ask him to become proficient in ferreting out the small defects of the product and criticize him when he quickly develops super skill.

We expect him to find every single defect in a fast moving assembly line when we theorize that 100% inspection is only 85% to 90% efficient.

We demand that he have the ability of a tool and die maker, but are unwilling to pay him the same wages. We insist that he be competent as an operator, but expect him to think differently than an operator.

He deals with foremen, engineers, industrial engineers, production control men, yet he is an hourly worker somewhere between the middle and the top of the pay scale, often with little or no authority.

We have given him no special training in diplomacy and tact, yet he should be the ultimate diplomat of the shop. He must be able to reject the work of his neighbor in such a way that they can ride home together as friends.

We expect him to be a paper work expert and criticize when he is not spending eight hours a day next to the machinery.

We insist that he be consistent from one day to the next and that he have infallible standards of judgment, but we are the ones who frequently ask him to change these standards.

What Can Be Done For The Inspector

These are the things an inspector is, and these are the things we expect him to be. Painted this way we gather the mental picture of an extremely confused man, and in many cases this is what we have. How do we end this confusion and have the inspector do the best job possible?

First, let us spend some time in the selection of the inspectors. Through a rather involved process York has developed a selection test which picks persons who have the temperament that we have found best suited to inspectors. We pick only people who have high school educations. We pick only people who have actual shop operating experience. Apparently, there is no substitute for the experience of being an ex-operator. We make sure the inspector is able to read blueprints well and most important, is able to understand complicated written instructions.

Second - Define His Job

Before we allow him to become an inspector, we make sure that in his own mind he understands the quality objective, the definition of quality control, and the definition of his job.

We tell him that he wears two hats. First, it is his job to determine good and bad. Second, and most important, it is his job to help prevent defects by the intelligent discovery of defect trends as early as he can. We equip him with the rudiments of statistics.

Third - Outline His Scope Of Authority And Responsibility

We tell him that he must be flexible without compromising. We show him that temporarily accepting a product with a minor defect is not compromising his ethics, but this does not remove his obligation from quickly correcting the situation. We make sure that he understands the limits of his authority, and that he uses them judiciously.

We point out to him the dangers of becoming super critical. Most of all, we stress to him the importance in dealing with facts and try to develop a humble appreciation of his power in the shop without making a monster of him in attitude.

We tell him his job really is to find the facts that prevent the defects.

Fourth - Periodically Retrain Him

We have come to realize that even the best inspector periodically needs reaffirmation of all the precepts and concepts by which he is expected to operate.

In those areas where there is a definite tendency to become super critical, we periodically have him inspect selected items and then discuss his determination of the quality as opposed to those of the Service Department, Sales Department and Engineering Department.

Fifth - Provide A Base For His Decisions

We have also found that he must have a reference for his normal decisions and a useable policy. For this we have developed an inspection and quality control manual that is well indexed and can be used quickly for matters of policy and procedure.

Placement Of The Quality Responsibility

In our industry there is still a cyclic nature to the business, and we have found that it is totally impractical to hire and fire large numbers of inspectors yearly. In some ways this has been a blessing, as we have had to realize that we can only maintain a core of inspectors and their job must be more of an audit than a job of covering every nook and cranny.

This has pointed out to us the urgency of quickly returning the basic responsibility for quality to the foreman and operator and putting the inspector in the position of a helper and a fact finder.

We have not de-emphasized the inspector's position, but we have made sure that he is regarded as a helper in emergencies and normally, the auditor and fact finder. There is always the problem of maintaining this understanding between the inspectors and the operators, and

it presents a delicate task of daily adjustment to the chief inspector and inspection supervisors.

It is something that will not maintain itself, but it takes constant and daily adjustment.

So far I have talked about the inspector in terms of the human failings and strength of the two-legged variety of inspector.

Mechanical Inspection

Our modern quality control concept has led us to build into our equipment the inability to make a mistake and the ability to inspect for itself. It has also made us insist that the engineers design products which cannot be built wrong and the industrial engineers provide tools which meet the same test.

This has gradually led to a reduction in the number of inspectors required. Obviously, the choice between building in automatic inspection and using inspectors is always an economic decision.

Cost Of Inspection

Gradually and inexorably there has been a trend toward a change in our old concept of the ratios of indirect cost to direct labor cost and the inspector is a part of this change. As our factories become more automatic and we use less direct labor people, we find ourselves being squeezed by the accountants because their old standards of measurement of the ratio of inspectors to direct labor men no longer hold true. Upon close examination, and honest examination, we find that we can build in more prevention and probably reduce our actual inspection force proportionate to the reduction in direct labor. This is not an economic axiom, but in our own experience we found that we were inclined to make excuses and say that we could not keep pace in the reduction of inspectors when we were competing with automation. The facts have proven us wrong, and with real effort we have been able to substantially reduce our inspection force. The reason for this is the increased use of control before the fact and a much closer working with the engineers and industrial engineers to build checks into our production equipment. But perhaps most important is the fact that we have better qualified inspectors on the job.

Inspection Utopia

I would consider it a utopian situation if we reached the place where we would not need to define "what is an inspector", because as a breed they had disappeared. I do this bit of philosophizing because no matter how we look at it, inspection and the inspector are necessary because of our failure to do a perfect job somewhere along the line. We can only be satisfied when these failings no longer exist and hence there is no longer a need.

I hope that by looking into the past and talking about what an inspector was and what he is today, you have gathered new ideas about what he should be in your industry tomorrow.

OPERATIONS RESEARCH IN BUSINESS AND INDUSTRY

Leonard W. Swanson
Arthur Andersen & Co.

INTRODUCTION

Definition

Operations Research (1)(2) or OR as it is often called is a relatively new field of endeavor. It has been in existence for approximately 15 years under that name although there are isolated examples of its use many years prior to that time. Briefly OR is the application of the scientific method to decision making. The quality control specialist is well aware of the importance of the scientific approach to certain problems of decision making. Quality control might well be included in the area of operations research but I do not intend to elaborate on that point.

Historical Background

Operations Research had its real beginning during World War II when the various military agencies employed scientifically trained personnel to assist in the improvement of military operations. These scientists were extremely successful in aiding commanding officers with the many complex problems that faced them. In arriving at solutions to these problems it was demonstrated that our wealth of scientific knowledge could be applied to many of the complex problems of management that had previously been handled by intuition or judgment. Since the war many OR groups have been formed to aid in the solution of management problems. These teams are used primarily in aiding executives by giving a quantitative evaluation of the alternative decisions that might be made--the executive however must still make the decisions.

Method of Attack

Typically, a problem is attacked by first defining the goal of the operation in a clear and concise manner. This includes the definition of the problem together with the measure of effectiveness to be used. Many times this portion of the problem requires more work and time than any of the other phases. The second stage involves the construction of the model which may or may not be mathematical in nature. This phase of the problem includes the collection of the pertinent data, which in itself is often time-consuming and costly. Following the collection of the data, a third phase consists of the solution of the problem together with the testing of the model for accuracy of representation. At this point it may be necessary to alter the model in the light of the testing that has taken place. If a new model results, then it must be retested. Once a satisfactory model has been constructed, and solutions have been obtained, it is necessary to present the results in terms of a workable operation. This final phase is extremely important since it presents management with the various alternative courses of action together with an evaluation of each.

Range of Problems

The range of problems (3) (4) being handled by Operations Research methods is extremely broad. The fields of application are in such widely scattered areas as administration, production, inventory, manufacturing, transportation and research. Your Administrative Applications Division of the American Society for Quality Control was organized "to promote the installation of statistical techniques in administrative fields." The opportunities for your using OR methods are essentially unlimited. In order to assist you in establishing the nature of Operations Research some pertinent examples are given in regard to several specific areas.

INVENTORY CONTROL

Inventory Costs

Inventory problems (5) arise because of the conflict of operational costs--these costs fall into the three main categories of Procurement, Possession and Shortage. In the area of procurement are all the costs associated with each order for replenishment. The total procurement cost depends upon the number of orders placed during a period of time. Possession costs consist of such items as warehousing, interest on capital, insurance and obsolescence. Investment in inventories can commit a considerable portion of a firm's capital--consequently savings from using modern methods of inventory control can be important. The third important cost associated with inventory problems is the cost of a shortage.

Measure of Effectiveness

Although the measure of effectiveness used in connection with inventories is not unique, one measure that is often used is the total variable operating cost. If such a measure of effectiveness is used, then the problem becomes one of determining the policy of handling the inventory so as to minimize the sum of procurement, possession and shortage costs. Since it is extremely difficult to assign a cost to the shortage of a particular item, the problem is often rephrased in terms of an allowable probability of an outage. Management is usually willing to specify a criterion in this area by specifying for example that it is permissible to run out of a particular item once in five years or whatever period of time is appropriate.

Economic Order Quantity

On the basis of establishing the quality of service by means of an allowable probability of an outage, it is then possible to establish an inventory management policy which will meet the service requirements and determine order point and order quantity so as to minimize the costs due to procurement and possession. For each inventory item at a storeroom, there is a variation in quantity balance. When the inventory reaches the level "K", called the reorder point, a quantity "Q", called the reorder quantity, is ordered. The lapse of time until the order is received (called procurement lead time) is extremely important in determining the reorder point. Formulas for economic order quantities are fairly well known and the rather standard

$$Q = \sqrt{\frac{2YS}{IC}}$$

is usually acceptable. In this formula:

Y = annual use of item in physical units
S = procurement cost per order
I = possession charge in per cent
C = unit cost of item

Reorder Point

The methods of determining reorder point "K" are not so well known however. Historical data on usage is of course extremely important. Such data is fit to an appropriate curve, usually a Poisson distribution (6). This usage pattern, together with management's decision on allowable outages, provides the necessary information for determining the reorder point.

Implementation

If means of automatic data processing are available, the computations can be handled quite easily. Such a computer operation would determine when the reorder point had been reached and then determine the quantity to be ordered. This would be done on an exception basis, printing the pertinent information on those items which need attention. If computing equipment is not available then charts or nomographs can be used. In any case, the reorder point determination is closely associated with the Poisson distribution as determined by the past usage of the item.

Value

This is a relatively simple method of inventory control --much more complicated ones have been developed. However, the attention that many firms give to the economic control of their inventory is so small, when the size of the investment is considered, that very real savings can be achieved. Several cases in our experience have shown cost reductions of hundreds of thousands of dollars per year for even moderate-sized firms.

INFORMATION PROBLEMS

General

Problems in this area arise when certain information is needed. There are both errors and costs associated with the process of data collection. The problem to be discussed here pertains to the railroad industry although a similar application has been made in other transportation fields.

Revenue Estimating

The problem here is to obtain an estimate of a railroad's full carload revenue. This represents revenue actually earned for a specified month with the stipulation that such information is to be available within the first few days after the close of the month's business. Such information cannot be obtained on an actual basis without going to extreme expense since the normal process is to wait for a

report from each of the final carriers. Such information is never available until the latter part of the month and sometimes not until 60 to 90 days later. The revenue estimate had not been a problem until recently, since the accounting reports were normally rendered about the 25th of the month. With the installation of a new accounting system such reports were to be made no later than the 10th of the month, and consequently it became necessary to devise a means of determining that revenue.

Scientific Sampling

Often many different railroads handle a particular car and the actual computation of the apportionment of the revenue to each of the railroads is quite a task. Historical information can be used to determine revenue per car but such figures vary so much from month to month that average revenue per car methods did not seem advisable. The labor involved in complete computation was so large that consideration was given to obtaining the estimates by taking a sample of the traffic on a random basis. Sample estimates are not only accurate enough for management planning but surprisingly are often more accurate than a 100% determination. Scientific sampling (7) takes into account the revenue characteristics of the various commodities hauled as well as the variations due to seasonal effects. Monthly revenue characterisitics are studied in order to assist in designing the sample.

Classification of Traffic

There are four classes of full carload traffic handled by the railroad. These are (a) local, handled completely on the line; (b) interline received, started by another railroad but finished on the line; (c) interline forwarded, started on the line but finished elsewhere; and (d) intermediate, started elsewhere and finished elsewhere but carried intermediately on the line.

Sampling Procedure

At the present time, the revenue for local traffic is the only one which can be obtained on a 100% basis without undue expense. All other classes of traffic would require excessively costly procedures to obtain complete revenue information and so it is in this area that the sample is chosen. On the basis of the accuracy required, a sample size of 1 in 15 was chosen for the nonlocal traffic. Worksheets have been drawn up on the basis of using 10 subsamples. Carloads are chosen from appropriate reporting forms by the use of random numbers and, for each car chosen, the revenue due the railroad is computed and entered on the worksheet. The totals for all apportionments are obtained at the end of the month and multiplied by fifteen in order to obtain an estimate of nonlocal revenue which, when added to the local revenue, gives an estimate of the total full car revenue. Adjustments to this figure are made on the basis of other reports in order to obtain the total operating revenue for the railroad.

Standard Error

The division into subsamples is made for the purpose of facilitating the computation of the standard error of the

estimate. The standard error on a cumulative basis continues to reduce throughout the year until the final figure is less than 0.5%. This means that there is 95% confidence that the error due to sampling is less than 1%.

Advantages

The procedure described for estimating revenue has several particular advantages. These are:

1. The cars are chosen from actual traffic and, hence, no historical information is needed. Unusual traffic characteristics will be found by the sample.
2. Changes in freight rates do not alter the procedure, except in computing the apportionments, at which time the new rate schedules are used.
3. Accuracy and corresponding risk of the estimate are determined precisely.
4. Revenue estimate is available by the 6th of the month.
5. Costly and inaccurate carry-over revenue computations are eliminated.
6. Railroad personnel can carry out the plan after it is set up.

Since the cost of making revenue apportionments is extremely high for any railroad, serious consideration should be given to using sampling for the actual settlement of interline accounts. The errors due to the sampling could quite likely be more than offset by the savings effected. Several airlines have already adopted this progressive move in accounting.

PRODUCTION ALLOCATION

Changing Market

The market for some manufactured goods changes rapidly, and consequently manufacturing facilities are not always in balance with the requirements for production. The purpose of the following example is to indicate a method of approach for the proper allocation of production to available manufacturing plants so as to maximize the profit over a specified period of time. The method may be used for either long-range or short-range planning. If the demands for each of a group of products are known and the requirements for producing these products are also known, then it is possible to determine the optimum allocation of production. As an illustration, consider the situation created by the reduced demand for a series of products which are presently being manufactured in several plants. On the basis of forecast demand for the products, the allocation of production can be determined so as to minimize the total variable cost.

Capacities and Demands

To be specific, consider a 7-plant operation. Suppose that the plants are allowed the option of operating on a 2-shift basis. Since the capacity for production is usually different on each of the two shifts, let these be designated as C_1, C_2,----------, C_{14}, where C_1, and C_2 are the first

and second shift capacities of the first plant and so on. The measure of capacity used is not unique but probably would be square footage, direct hours of labor or something similar. Suppose there are 100 products being manufactured. Let D_1, D_2, D_3,----------, D_{100} be the demands for each of these products. It is important that these demands be given in the same units as the capacities of the plants.

Cost Matrix

It is first necessary to determine the cost associated with producing each of the products at each of the plants. In order to reduce the computation, only the costs which will vary with a change in location of the production will be used. Thus the cost of the material for making the product may not enter as a variable cost, since the cost of the material will usually be essentially the same at each plant. The items of cost which are likely to enter as variable are the cost of getting raw materials to the plant, gas, electricity, taxes, material movement in the plant, cost of direct labor if rates vary between plants, distribution cost of final product and possibly several others. The task of gathering such information is sometimes sizeable. The costs when gathered are then arranged into a variable cost matrix as shown

COST MATRIX

	Plant 1	Plant 2	------	Plant 14	Demands
Product 1	$A_1,1$	$A_1,2$	------	$A_1,14$	D_1
Product 2	$A_2,1$	$A_2,2$	------	$A_2,14$	D_2
Product 3	"	"	------	"	"
"	"	"	------	"	"
"	"	"	------	"	"
"	"	"	------	"	"
Product 100	$A_{100},1$	$A_{100},2$	------	$A_{100},14$	D_{100}
Capacities	C_1	C_2	------	C_{14}	

Optimum Allocation

The completed matrix presents the variable cost for manufacturing each product at each plant. Up to this point it has been assumed that each product could be produced at each plant--provisions will be made later for handling various restrictions on production. If, now, one uses the notation that Pij represents the amount of product i to be allocated to plant j, the total variable cost of production is given by

$$\text{Cost} = A_{1,1}P_{1,1}+A_{1,2}P_{1,2}+\text{----------}A_{100,14}P_{100,14}.$$

The specific goal here is to allocate production in such a fashion as to minimize the cost given above while satisfying the demands exactly and yet not exceeding the capacities of any of the plants. The demands for each of the products have been added as a right-hand column in the cost matrix and the

capacities are given as the bottom line in the same matrix.

Linear Programming

Since it is assumed that the variable cost of producing a given product at a given plant varies directly as the amount produced, this problem becomes a linear programming (8) one. The assumption is essentially correct unless the amounts allocated to a given plant fall below a 70% usage of that plant. Furthermore, if the total of all of the demands is exactly equal to the total of all the capacities, the problem reduces to one which can be solved by the "transportation" method. Actually a likely situation is that the total demand will be less than the total capacity and consequently it is not a true transportation problem. Because of the reduced computation in the transportation type problem a fictitious product is added in order to make the problem come under that classification.

Computer Computations

The computations to determine the optimum solution can be carried out very nicely on an intermediate-size computer, and a solution has been obtained for a similar-size problem in approximately an hour's machine time. The final solution must be examined carefully to make sure that it is realistic. In case adjustment must be made to the optimal solution, the additional costs due to such adjustments can be determined quite easily.

Variations on Problem

There are several ramifications to this problem which should be pointed out. First of all, the above discussion has been based upon the assumption that there is no variation in fixed costs. If demand should drop considerably there is the possibility of closing a plant--which in turn would lead to a reduction in fixed costs. If one then allows a plant to be closed the fixed costs become variable and any savings in this area must be added to the savings in what has previously been termed the total variable cost, to determine the ultimate savings. If costs are associated with the shifting of production from one plant to another, these costs must be applied against the savings. In some cases it is possible to incorporate the costs of shifting production in the cost matrix itself.

Restrictions

In the cost matrix, certain artificial restrictions can be incorporated. For example, should management indicate that it is impossible to produce a given item at a given plant, that particular cost can be assigned arbitrarily high in order to prevent that allocation of production. A similar assignment can be made to the costs associated with each product at a particular plant, particularly if management desires to see the effect of closing a particular plant. Proper design of the computing procedure also makes it possible to change all the costs associated with a particular product, should such be desired. Such a design also makes it possible to evaluate either the addition of new products for long-range planning or

the elimination of the production of certain products from the line.

Savings

The procedure described above has actually been carried out in several cases. One of these resulted in an annual savings in excess of one million dollars. The procedure is flexible enough that it can be carried out quite easily providing the appropriate data is available.

REFERENCES

(1) Operations Research for Management, Volumes I and II, Baltimore, Maryland, Johns Hopkins Press, 1954.

(2) Operations Research, Journal of ORSA, Baltimore, Maryland, Operations Research Society of America, 1953-1956.

(3) Management Science, Journal of the Institute of Management Sciences, Baltimore, Maryland, Institute of Management Sciences, 1954-1956.

(4) Operations Research, A Basic Approach, Special Report No. 13, New York City, American Management Association, 1956.

(5) WHITIN, T. M., The Theory of Inventory Management, Princeton, New Jersey, Princeton University Press, 1953.

(6) FELLER, WILLIAM, An Introduction to Probability Theory and Its Applications, New York City, John Wiley & Sons, 1950.

(7) DEMING, W.E., Some Theory of Sampling, New York City, John Wiley & Sons, 1950.

(8) CHARNES, A., COOPER, W. W. and HENDERSON, A., An Introduction to Linear Programming, New York, John Wiley & Sons, 1953.

FRACTIONAL REPLICATION IN INDUSTRIAL EXPERIMENTATION

Cuthbert Daniel
New York City

I propose to talk on two aspects of fractional replication in the 2^{p-q} series: first on sequences of fractional replicates and then on improved error-estimation. Each of the two proposals made below is aimed at redressing, partially at least, one of the outstanding defects of fractional replication as usually practised.

The first defect is that of size. If eight factors and all their two factor interactions must be estimated, 64 runs (treatment combinations) must be made. If full randomization is enforced, no conclusions can be drawn until all runs are completed. The second defect is the lack of a proper error term in the standard designs.

Sequences of fractional replicates

Let us suppose that it is desired to obtain estimates of the first-order effects of p factors, independently variable, and of their $p(p-1)/2$ two factor interactions. There are, then, $p(p+1)/2$ mutually unconfounded estimates required. For $p = 6,8,11$, the minimum numbers of runs required are 32, 64 and 128. But it often happens that, after one of these large sets has been completed, most of the factors varied and most of the two factor interactions appear to be of small influence relative to the others. The experimenter may feel that he could have found the bigger effects in fewer runs; that he had to wait too long to discover the major effects. If only a part of the whole design could have been completed first, he might have been able to say either:

a. Factors A, D and E show that they are far and away the most influential in the present range of their variation. I can draw my conclusions now. OR
b. Factors A, D and E appear most influential but I must have better estimates and I must know more about their interactions in the range of variation under study. OR
c. Changing A made a big improvement. I want to increase it still more. There is no point in continuing experimentation with the present range. OR
d. Nothing much shows up, but my experimental error looks surprisingly large. We'll have to do another round to increase our sensitivity.

Consider, as we mathematical statisticians say, the half-replicate of the 2^6 (called here a 2^{6-1}) requiring 32 runs which gives all six first-order effects and all 15 two factor interactions (abbreviated 2 f.i. from now on) mutually unconfounded. Of course since it is a half replicate each of these must be confounded with something, and the cryptic rubric

I + ABCDEF

shows what is measured along with what. As examples: the main effect of varying C is measured along with the five factor interaction ABDEF; the 2 f.i. DE is measured along with ABCF. No smaller factorial experiment will measure all 15 2 f.i. and all 6 main effects mutually unconfounded, but there is a set of 16 runs, half of the above half rep. and therefore designated 2^{6-1-1}, that will permit estimates of all the main effects. Of course this set of 16 runs, being a quarter rep. produces contrasts each of which estimates the sum of four effects. The best that can be done is to permit each main effect to be measured along with <u>one</u> 2 f.i. The confounding pattern is given by

I - ABC - DEF + ABCDEF.

Thus the "A contrast" measures A - BC - ADEF + BCDEF. The 16 runs are specified by the four "generators", <u>ab</u>, <u>bc</u>, <u>de</u> and <u>ef</u>. Multiplying these together in all combinations and adding the run designated (1)- meaning that with all factors at their low levels- gives the 16 runs required.

If after analysis the experimenter opts for completing the 2 f.i.-clear design, he has only to do the 16 runs derived by multiplying the set already specified by <u>ad</u>.

Similarly for the other favorable but large fractional replicates in the 2^{p-q} series. There is for each a most favorable (or least unfavorable) sub-fraction, 2^{p-q-r}, that will separate and detect large main effects. For p = 8 the smallest 2 f.i. clear plan is the 2^{8-2}. The eight main effects can be separated from each other, but not from the 2 f.i. interactions, in 16 runs, thus in a 2^{8-2-2}.

A confounding pattern for the 2^{8-2-2} is given by:

I - ABC - CDE - EFG - GHA

+ ABDE + AEFH + BCGH + BDFH + CDFG

- ABDFG - <u>ACDFH</u>- BCEFH- <u>BDEGH</u>

+<u>ABCEFG</u> + ACDEGH.

The four 3 f.i. can be taken as generators. The underlined interactions are viewed as the "permanent members". These three give the confounding pattern of the 2 f.i. clear 2^{8-2} design that the experimenter may have to complete.

The first 16 runs may be generated by <u>hab</u>, <u>bcd</u>, <u>def</u>, <u>fgh</u> If another set of 16 runs can be done, then three of the 3 f.i. can be removed from the confounding pattern. For example -ABC, -CDE, and -EFG are removed if the second set of 16 runs is generated by <u>be</u> times each member of the first set The multiplier must be odd with respect to the three removed

3 f.i. but even with respect to the underlined members. After doing the first 16 runs, we are not committed only to the 2^{8-2} specified by the underlined three interactions. The three not underlined (two fives and a six) give us the only alternative. Thus if it was decided that -CDE, -EFG and -GHA should be eliminated, then the new 16 runs are generated by dg times the old set, but this puts us definitely in the 2^{8-2} defined by the 5 and 6 f.i. not underlined.

Partially duplicated fractional replicates

"In certain situations" as we say when we cannot specify them exactly, it is important that a fairly unbiassed estimate of the run-to-run error be obtained. The inveterate habit of pooling higher-order interactions tends to spread to the 2 f.i. since there appears no compelling reason why the latter should be in a qualitatively different category than the 3 f.i. Of course whenever a 2 or a 3 f.i. looks big, we quietly snatch it from the pool and test its "significance".

A natural compromise between this disreputable practice and the expensive blamelessness of full replication is that of partial duplication. Some runs, then, will be repeated but not all. It only remains to decide which, and to show the consequences in analysis and in interpretation.

To take a simple example first, we choose a 2^3 with half its runs replicated. We will designate this as a $2^3 + 2^{3-1}$. Use I - ABC for the half-rep.;call the usual difference between the sum of all results at high A and the sum of all results at low A (A) ; and let the symbol $\hat{A}$ stand for the estimated average A-effect.

By standard least squares operations, or by inspection, we find pairs of normal equations like these:

$$(A) = 6\hat{A} - 2\widehat{BC}$$

$$(BC) = -2\hat{A} + 6\widehat{BC} .$$

From these we find:

$$16\hat{A} = 3(A) + (BC)$$

$$16\widehat{BC} = (A) + 3(BC)$$

Also the variance of each of the estimated effects is 3/8 of the run-to-run error variance, and the correlation between the pairs that appear together is -1/5. Those that appear in separate pairs are uncorrelated. Thus similar equations and similar relations hold for the pairs $\hat{B}$ and $\widehat{AC}$, and $\hat{C}$ and $\widehat{AB}$. The among-run error variance is estimated with only four degrees of freedom in this example, by 1/8 the sum of squares of differences between the four duplicate pairs of runs.

A more serious example is given by the half-replicated 2^{5-1}. The full half-rep. may be assigned the confounding pattern I - ABCDE. The duplicating quarter-rep. is given the pattern I - ABC + DE - ABCDE. The duplicated runs are those generated by ab, ac, and de. The unduplicated runs are found by adding the generator ad.

Two sorts of pairs of normal equations emerge.

$$(A) = 12\,\hat{A} - 4\hat{BC} \qquad \text{and} \qquad (D) = 12\,\hat{D} + 4\,\hat{E}$$
$$(BC) = -4\,\hat{A} + 12\hat{BC} \qquad\qquad (E) = \hat{4}\,D + 12\,\hat{E}$$

Their solutions look like this:

$$32\hat{A} = 3(A) + (BC) \qquad \text{and} \qquad 32\hat{D} = 3(D) - (E)$$
$$32\hat{BC} = (A) + 3(BC) \qquad\qquad 32\hat{E} = -(D) + 3(E).$$

I am indebted to Miss Edith Reid who showed me the useful shortcut of carrying through the usual Yates algorithm using the sums of pairs where they occur, to obtain the values of all the contrasts simultaneously

The 2^{8-2} can have a 2^{8-2-2} in 16 runs appended for error estimation in the same way. For the full 2^{8-2} the best confounding pattern is given by

$$Q_1 = I - ACDFH - BDEGH + ABCEFG.$$

For the duplicated runs, use

$$Q_2 = (\, I - ABC - CDE \,) \times Q_1 .$$

Four types of groups of normal equations appear. Their solutions are also given.
I. For $\hat{A},\hat{BC},\hat{GH}$; $\hat{C},\hat{DE},\hat{AB}$; $\hat{E},\hat{FG},\hat{CD}$; and $\hat{G},\hat{AH},\hat{EF}$.

	$8\hat{A}$	$8\hat{BC}$	$8\hat{GH}$	(A)	(BC)	(GH)
1)	5	-1	-1	1	0	0
2)	-1	5	-1	0	1	0
3)	-1	-1	5	0	0	1
4)	28	0	0	6	1	1
5)	0	28	0	1	6	-1
6)	0	0	28	1	-1	6

II. For $\hat{AE}, \hat{BD}, \hat{FH}$; and $\hat{BH}, \hat{CG}, \hat{DF}$.

	$8\hat{AE}$	$8\hat{BD}$	$8\hat{FH}$	(AE)	(BC)	(FH)
7)	5	1	1	1	0	0
8)	1	5	1	0	1	0
9)	1	1	5	0	0	1
10)	28	0	0	6	-1	-1
11)	0	28	0	-1	6	-1
12)	0	0	28	-1	-1	6

III. For $\hat{B}, \hat{AC}$; $\hat{D}, \hat{CE}$; $\hat{F}, \hat{EG}$; and $\hat{H}, \hat{AG}$.

	$8\hat{B}$	$8\hat{AC}$	(B)	(AC)
13)	5	-1	1	0
14)	-1	5	0	1
15)	24	0	5	1
16)	0	24	1	5

IV. For $\hat{AD}, \hat{BE}$; $\hat{AF}, \hat{EH}$; $\hat{BF}, \hat{DH}$; $\hat{BG}, \hat{CH}$; and $\hat{CF}, \hat{DG}$.

	$8\hat{AD}$	$8\hat{BE}$	(AD)	(BE)
17)	5	1	1	0
18)	1	5	0	1
19)	24	0	5	-1
20)	0	24	-1	5

Summary

1. A method of ordering parts of fractional replicates in the 2^{p-q} series has been outlined so that large effects can be detected or ruled out before the full set of runs is completed.

2. The design and analysis of partially duplicated fractional replicates has been sketched. The purpose of partial duplication is to secure a less biassed error estimate. The disadvantages are that a slightly greater computational effort is required, that small intercorrelations between some of the estimates are present, and that a slightly larger variance per run is obtained.

3. The combination of these two compromises is in preparation.

QUALITY CONTROL'S OBLIGATION TO MANAGEMENT AND CUSTOMER ON RELIABILITY OF COMPLEX WEAPONS

BY
ALMERON BEDFORD
REDSTONE ARSENAL

Gentlemen, I speak here today as an individual and not in any official capacity as a Government employee.

There are a lot of big words in the subject of today's discussion. Big words lead to big definitions and then the fun begins. It's like the three Frenchmen trying to explain to an Englishman the inability of the wife of one of them to have a baby. The husband said she was "unbearable;" the second corrected him and said she was "inconceivable;" and the third tried to improve on the others with his definition- the woman was "impregnable!"

And so it is with Quality Control - Quality Assurance, Reliability, Accuracy, Satisfaction Guaranteed, or what have you. I'm sure we all have the same basic idea in our minds, but we clothe it with different connotations in accordance with the usage dictated by our past experience or the school of thought with which we are most familiar.

Webster only contributes to the confusion by giving a dozen different definitions of "Quality." When we use it in the field of Guided Missiles, we prefer to think of Quality in terms of "DISTINCTIVE CHARACTERISTICS." - Characteristics that can be recognized, defined and measured so that they can be specified and designed into the product and be sufficiently identifiable to make it possible to determine later whether or not they are actually in the finished product.

Quality - as we see it - is the key to successful performance - it must, by its very nature, be specifically defined to the last distinctive characteristic. It's either right or wrong - there can be no compromise.

So, if we can get this picture of Quality as a set of definable, recognizable, measureable - distinctive characteristics - then we can start talking about our two basic responsibilities: how to get this required Quality into our Guided Missiles in the first place - and then how to reassure ourselves that it is actually there in the finished product, in the second place.

So that you will better understand our problem, tho, let's digress for a moment to distinguish between Quality Control as practiced generally in industry and as we visualize its function in the Guided Missile field.

In industry, the guiding factor in the implementation of the Quality Control concept is the competitive market that must be considered in the battle for economic survival. The product will not necessarily be the best, or the worst, that can be made, but must be of a Quality Level - somewhere between the two - which will appeal to the greatest number of potential customers - at a price they are willing to pay - and at a manufacturing cost that will allow for a reasonable profit. In arriving at this level of quality, industry frequently can also take a calculated risk. If manufacturing costs can be reduced at the expense of causing one failure in ten, but at a savings in production costs in excess of the cost of one unit - they can afford to replace the defective unit free of charge and still make more money than on the original operation. But, we can't take risks of any kind - calculated or otherwise. The ultimate customer is the enemy and if we deliver a defective unit at a critical time, there will be no chance to replace it "free-of-charge." It may mean the difference between a missile hitting or not hitting some vital installation. We have just one consideration - or problem - and granted it's a big one. We have to produce the most effective and reliable product that the tactical situation calls for - within the limits of human capabilities - and regardless of cost. Within reason, of course, but who is going to place a price tag on survival, beyond which we will not go? If an iron rivet will do the job - certainly we will specify and use iron rivets. But, if we can only design missiles that will only function properly with solid gold rivets - then, by golly - we'd better use gold rivets if necessary, until Fort Knox goes dry.

Now, of course, we are both customer and manufacturer, and while the former considerably outweighs the latter, we do have an understanding of both sides of the fence. In addition - even where we are the customer - because of the fact that Guided Missiles are so vital to our very continued existence - the Government has assumed the prerogative of going into the manufacturer's plant and exercising the function of Controlling Quality. It's accomplishment is either directly by the Government's own personnel or indirectly by Government surveillance of the Contractor's Quality Control personnel. It is primarily an inspection function, but because of factors which go beyond routine inspection in the course of assuring ourselves that the required quality is really there, we have defined this function as Quality Assurance. It is not to be confused with Quality Control which provides for the original inclusion of the required quality in the design of the product, but merely reassures us that everything has gone according to plan - or if not, of course, identifies the trouble spots so that necessary remedial measures can be instituted at once.

That brings up the question - where does Quality Control begin and where does it end. In its broadest sense, we believe it starts

with the first line on the drawing board and ends when the missile hits the target. It means taking every possible, conveivable, reasonable step to make sure that the end product is as effective, reliable, accurate, dependable and efficient in terms of the mission to be accomplished that it is humanly possible to produce. Sure, we need qualified quarterbacks to call the signals - sort of trouble-shooting efficiency experts - but, Quality "achievement" is everybody's business and the cost is eternal vigilance for everyone connected with the business of supplying materiel to the Armed Forces.

We reach a point here of divided opinion. Many differentiate between Quality Control as we have defined it and Reliability. Basically, this is a fallacious assumption, for if the product has been properly conceived, designed, manufactured and handled it must be reliable or else the original Quality specifications are wrong. If it hasn't the right quality designed and built into it in the first place, it never will be reliable. Now, if compromises have been effected with the original Quality specifications, if specified procedures have not been followed, or any of a hundred other possible abuses of the system as originally conceived have occured, then, of course, reliability is affected. It all goes back to our premise that the product must have the right Quality specifications to begin with - granted there will be trial and error before the final specifications are crystalized. You can put corn syrup in the gas tank of a Cadillac and ruin the engine. Reliability has been destroyed, but up to that point the desired quality and the potential reliability were both present. In other words, reliability could be defined as the prevention or elimination of all possible abuse or misapplication of the original quality specifications. But, again it all boils down to the basic objective and regardless of what you call it, you begin to visualize the magnitude of the problem that faces our engineers, our Specification Writers, etc.; first in determining what the correct qualities are - specifying them, and then seeing that they get into the product. It would be easy enough to throw discretion to the winds - as is often done in war time or cases of extreme national emergency - but, they are no more anxious to destroy the solvency of the Government in peace time than they are to see our country improperly defended in war time.

Where we are convinced it will not affect the efficient functioning of the missile, or any part, we do specify certain tolerances which lie within the specified qualities. And where it is, by the same token, possible to specify the industrial conception of a "low quality" product (the iron rivet against the gold rivet) we do just that. But, when we do approach these limits of acceptability as closely as we dare, we must then insist on exact compliance. They have a phrase for it at the Frankford Arsenal Military Specification School - Minimum Requirements (in quality) but, Maximum Severity (in meeting those requirements).

There's a psychological approach here, already covered briefly - one that perhaps lies in the field of Human Engineering. Its our feeling that Quality Assurance, lets call it - should be everybody's responsibility - not just that of a group of specialized trouble shooting, efficiency experts. It's too big a problem to solve with regulations and supervision alone. The human element is too paramount in every phase of the operation and without the personal cooperation of everyone connected with the program, the best laid plans can conceivably go haywire. Granted every step should be taken - people being only human - to eliminate the human element as far as possible, but we still have to consider it in all its potential notifications.

Industry can afford to be opportunists - we must be perfectionists.

ORGANIZING FOR AIRCRAFT QUALITY CONTROL

Herbert B. Epstein
Chance Vought Aircraft, Incorporated

Introduction

In order to describe to you gentlemen the scope of Quality Control as practised in the Aircraft Industry I have arranged for your review a series of illustrations which show in general chronological order the development of a Quality Control organization and system designed to keep pace with the ever increasing complexity of aircraft and guided missiles.

Although this experience is principally Chance Vought Aircraft experience it may be noticed in the evolution of the Quality Control function that the considerations for evolving this function are fundamental to the entire field of Quality Control.

Evolution of Quality Control Function

Inspection 1946 - From the inception of Chance Vought Aircraft to 1946 the Quality Control function consisted principally of inspecting aircraft from the raw material phase through to the completed airplane. However, in 1946 it was recognized by Chance Vought top management that a need existed to plan for and monitor the increasing complexity of military aircraft. This was accomplished by a complete reorganization in 1946.

Process Control and Fabrication Control 1947 - Analysis of the prime areas of control requiring high technical competency resulted in the transfer of Engineering Department functions to Quality Control. As a result engineers were transferred from Engineering and hired to Quality Control from the outside to form the Process Control and Fabrication Control Sections in 1947. During this period the F6U-1 and F7U-1 were in their initial stages of fabrication.

Quality Electronics 1951 - With the introduction of the Regulus I Guided Missile it became apparent that a major effort had to be made to assure top quality electronic components and systems for the Regulus I program. To accomplish this the Quality Electronics Group was established under the Fabrication Control Section. This group was responsible for establishing preventive and corrective action for inspecting, calibrating and testing all electronic equipment on Missile and aircraft.

Quality Control Organization

From this evolution of the inspection function through to the plant wide concept of Quality Control ranging from design considerations through to customer acceptance, the Quality Control function developed. It consists essentially of two functions; an engineering function and an inspection function. The engineering responsibility includes working closely with the design engineers, planning new processes, and preparing the necessary instructions to monitor these processes in the shop in order to prevent problems; and trouble-shooting problems that do occur and take corrective action as necessary.

The Inspection Sections have the prime responsibility of assuring that all parts and assemblies accepted meet drawing and specification requirements through actual physical inspection and provides the necessary policing to see that the control procedures established by the Quality Control Engineering Sections are carried out.

Shop and Airport Inspection Sections - Illustrated is a flow diagram of inspection processing of raw material through to complete aircraft check-out and customer acceptance.

Process Control Section - The Process Control Section plans for and monitors all chemical, metallurgical and bonding processes, establishes necessary statistical quality control procedures on a plant wide basis, and controls all outside quality control activities.

Fabrication Control Section - The Fabrication Control Section plans for and monitors all mechanical processes, final assemblies, field operations and electronic operations.

Quality Control Program

It may be seen from the Quality Control organizational set-up that a complete and systematic quality control program for all models has been established. Functionally this program starts with the original design where Quality Control Engineers with production experience may confer with their design colleagues on new models to prevent problems that history has indicated may be expected. From this design phase through to analysing of customer and service complaints the feed-back cycle is completed.

Quality Control of Manufacturing Areas

In order to implement this program it is necessary to know where the major poor quality areas are so that manufacturing cost can be reduced by reduced scrap rates and production schedules may be met with more assurance. As an example of the method used by Quality Control to maintain and improve the quality of Chance Vought products, each Manufacturing area is reviewed for poor quality trends and repetitive discrepancies and malfunctions using the latest Statistical Quality Control methods. Just as the Process Control and Fabrication Control Engineers handle problems on the spot to get them corrected, the Statistical Quality Control Engineer objectively reviews by major area and as necessary specific unit the effectiveness of the detail Manufacturing and Quality Control effort. Through high point reports and other detail reports to all levels of management a constant quantitive level of quality control performance is maintained which is specific enough to help the line supervision improve the quality of their work. Through the use of statistical control methods which define normal variations in quality, deviations from a controlled operation are investigated and the elements contributing to poor quality are isolated and eliminated.

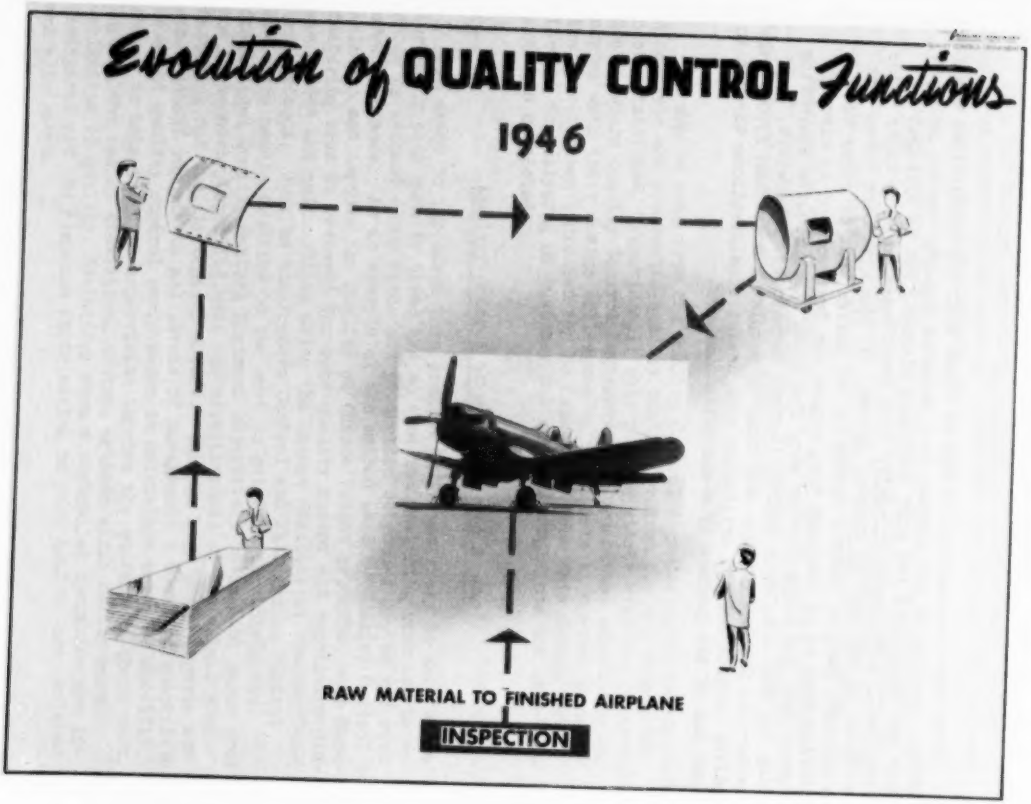
Statistical Quality Control Function

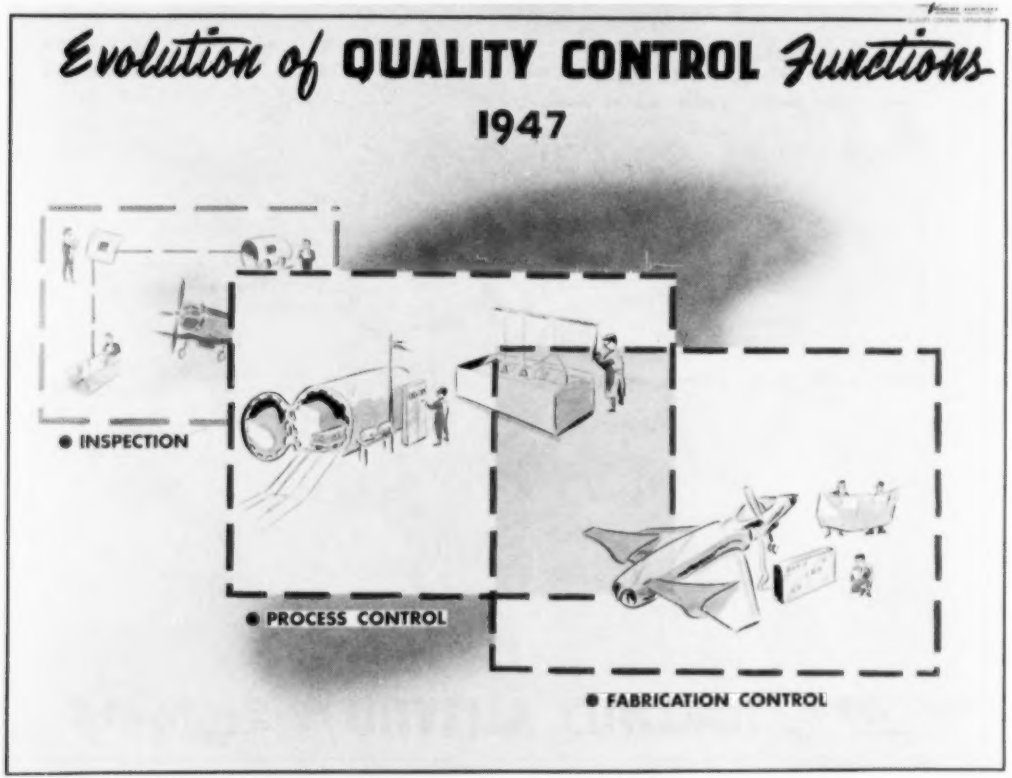
In order to keep the overhead costs down to a minimum and yet permit plant wide surveillance in sufficient detail to permit corrective action, the Statistical Quality Control Group has for the past several years been utilizing and further refining their IBM methods of opera-

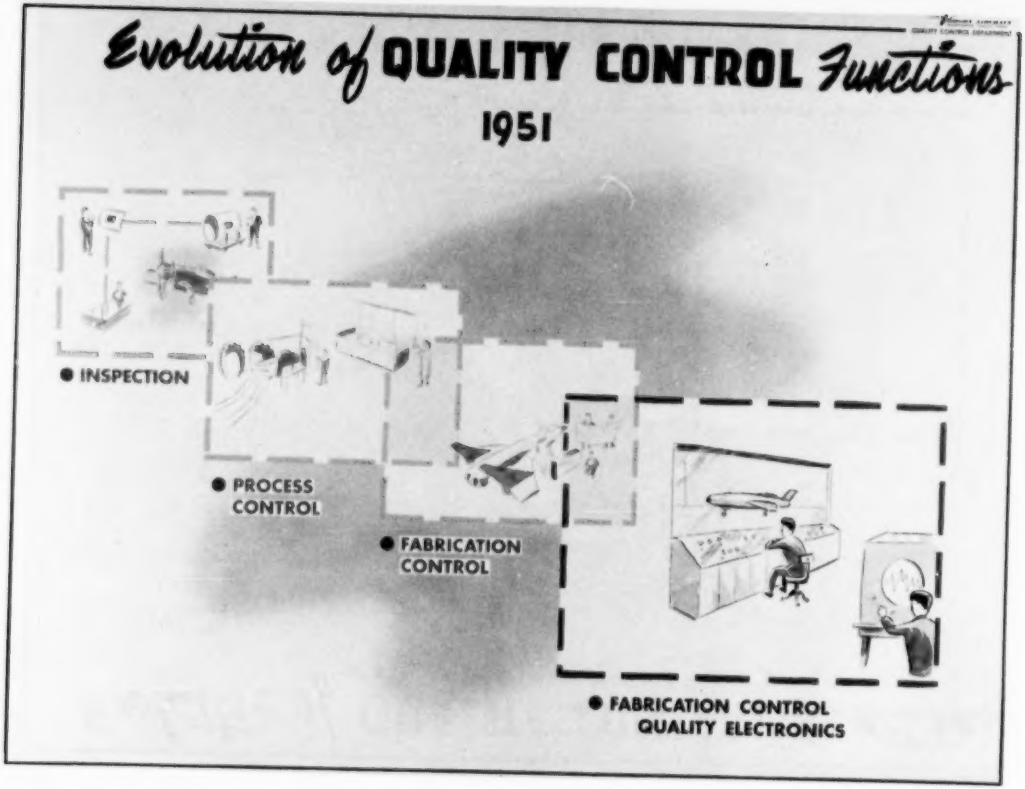
tion from the detail fabrication stage through to service complaints. Although this chart illustrates use of IBM methods in the Material Review system, it typifies the system used for controlling service complaints, vendor troubles, Government Furnished Equipment troubles, as well as all factors involved in the Manufacturing areas.

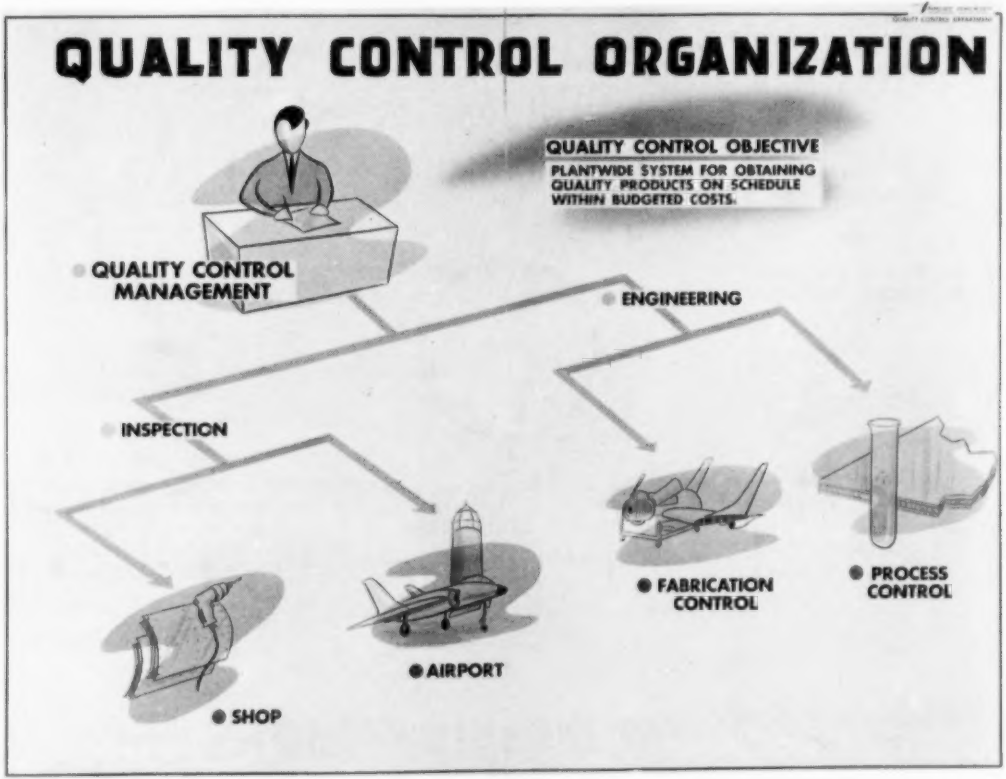
Conclusions

From this presentation it may be noted that the Statistical Quality Control function is one element in a balanced Quality Control program. It is not considered the most important element, but is arranged functionally and organizationally at a level comparable to other major functions in the Quality Control system. It has also been shown that an Aircraft Quality Control System is not unlike that of other industries when it is realized that within the framework of an aircraft company many job shops perform many functions which require monitoring using the best available Quality Control techniques. Therefore, whether the Company employs less than 50 people or more than 10,000, the Quality Control concept originates with design considerations through to final use of the product by a satisfied customer.


Evolution of QUALITY CONTROL Functions
1946
RAW MATERIAL TO FINISHED AIRPLANE
INSPECTION


Evolution of QUALITY CONTROL Functions
1947
● INSPECTION
● PROCESS CONTROL
● FABRICATION CONTROL


Evolution of QUALITY CONTROL Functions
1951
● INSPECTION
● PROCESS CONTROL
● FABRICATION CONTROL
● FABRICATION CONTROL QUALITY ELECTRONICS


QUALITY CONTROL ORGANIZATION
QUALITY CONTROL OBJECTIVE
PLANTWIDE SYSTEM FOR OBTAINING QUALITY PRODUCTS ON SCHEDULE WITHIN BUDGETED COSTS.
QUALITY CONTROL MANAGEMENT
ENGINEERING
INSPECTION
FABRICATION CONTROL
PROCESS CONTROL
AIRPORT
SHOP

QUALITY CONTROL FUNCTIONS
SHOP and AIRPORT Inspection Sections
• RAW MATERIAL & PURCHASED PARTS
• TEST LABORATORIES
• DETAIL INSPECTION
• ASSEMBLY INSPECTION
• FINAL INSPECTION
• SYSTEMS CHECKOUT & AIRPLANE SHAKEDOWN
• FLIGHT LINE INSPECTION
• CVA ACCEPTANCE
• CUSTOMER ACCEPTANCE
• DELIVERY
U
380

QUALITY CONTROL FUNCTIONS

PROCESS CONTROL *Section*

ENGINEERING

- CHEMICAL
  - INSTALL PROCEDURES TO PREVENT DEFECTS
- METALLURGICAL
  - CORRECTIVE ACTION TO PREVENT RECURRING DISCREPANCIES
- BONDING

QUALITY CONTROL TECHNICAL SERVICES

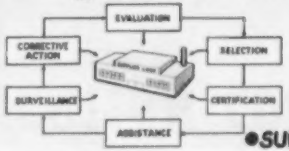
- STATISTICAL QUALITY CONTROL
  - ESTABLISH PROCEDURES FOR RECORDING, CONTROLLING & HIGHLIGHTING POOR QUALITY AREAS FROM DETAIL MANUFACTURE THROUGH FLIGHT TEST, DELIVERY AND SERVICE

- MATERIAL REVIEW, DISPOSITION AND CORRECTIVE ACTION

- STANDARDS

OUTSIDE QUALITY CONTROL

- VENDOR QUALITY CONTROL



- SUBCONTRACTOR QUALITY CONTROL

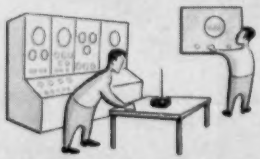
QUALITY CONTROL FUNCTIONS

FABRICATION CONTROL *Section*

- INSTALL PROCEDURES TO PREVENT DEFECTS

- QUALITY ENGINEERING

- CORRECTIVE ACTION TO PREVENT RECURRING DISCREPANCIES

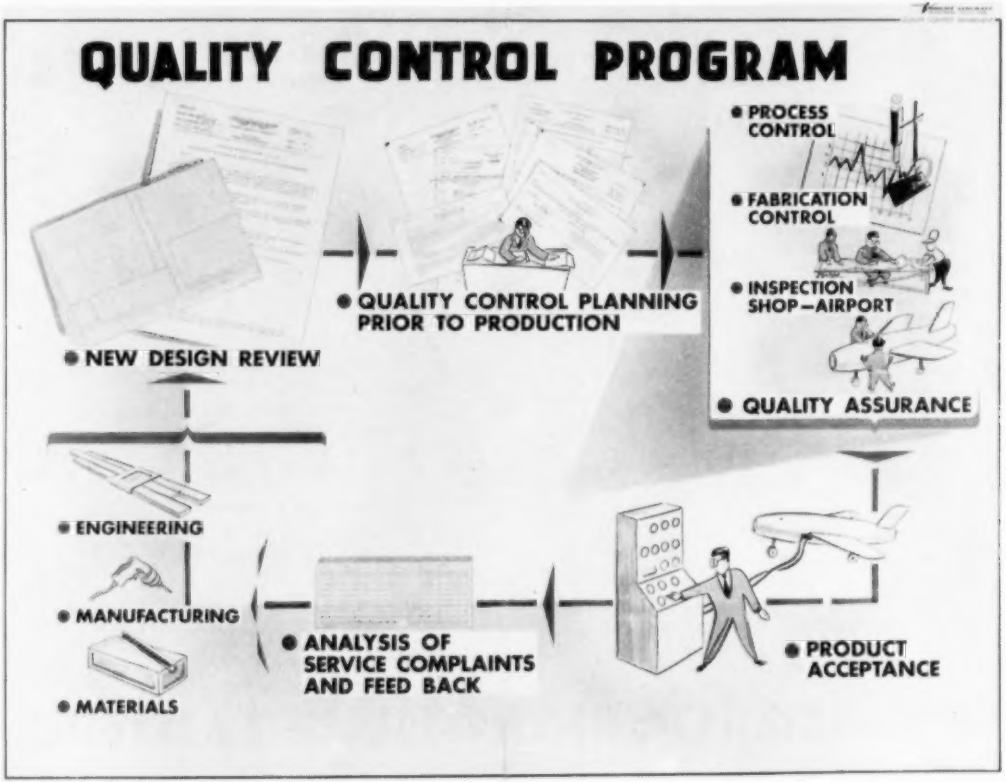


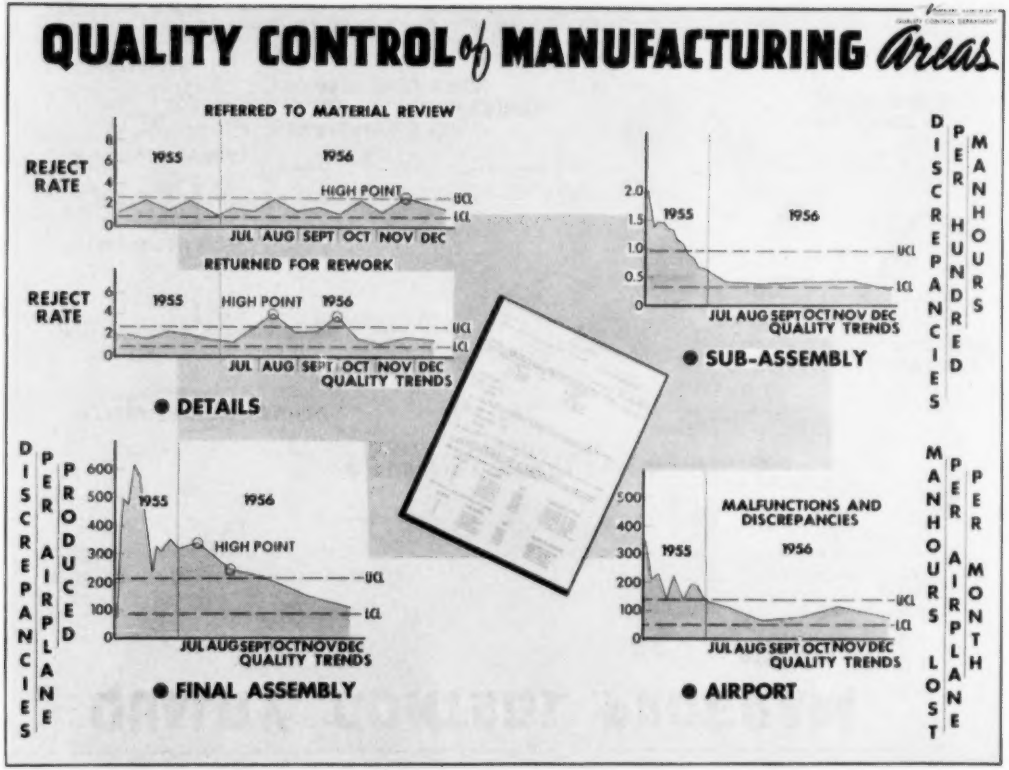
- RECEIVING INSPECTION OF ELECTRONIC COMPONENTS TO PREVENT MALFUNCTIONS

- QUALITY ELECTRONICS

- CORRECTIVE ACTION TO PREVENT RECURRING DISCREPANCIES

QUALITY ASSURANCE


QUALITY CONTROL PROGRAM
• PROCESS CONTROL
• FABRICATION CONTROL
• INSPECTION SHOP—AIRPORT
• QUALITY ASSURANCE
• QUALITY CONTROL PLANNING PRIOR TO PRODUCTION
• NEW DESIGN REVIEW
• ENGINEERING
• MANUFACTURING
• MATERIALS
• ANALYSIS OF SERVICE COMPLAINTS AND FEED BACK
• PRODUCT ACCEPTANCE


QUALITY CONTROL of MANUFACTURING Areas
REFERRED TO MATERIAL REVIEW
REJECT RATE
1955
1956
HIGH POINT
UCL
LCL
JUL AUG SEPT OCT NOV DEC
RETURNED FOR REWORK
REJECT RATE
1955
HIGH POINT
1956
UCL
LCL
JUL AUG SEPT OCT NOV DEC
QUALITY TRENDS
● DETAILS
DISCREPANCIES PER HUNDRED MANHOURS
1955
1956
UCL
LCL
JUL AUG SEPT OCT NOV DEC
QUALITY TRENDS
● SUB-ASSEMBLY
DISCREPANCIES PER AIRPLANE PRODUCED
1955
1956
HIGH POINT
UCL
LCL
JUL AUG SEPT OCT NOV DEC
QUALITY TRENDS
● FINAL ASSEMBLY
MANHOURS LOST PER AIRPLANE PER MONTH
MALFUNCTIONS AND DISCREPANCIES
1955
1956
UCL
LCL
JUL AUG SEPT OCT NOV DEC
QUALITY TRENDS
● AIRPORT

STATISTICAL QUALITY CONTROL Function
REJECTION
CODING
REJECT
REJECTED MATERIAL REPORT
PUNCH CARD CODE SHEET
KEY PUNCHING AND VERIFYING
MANUFACTURING
VENDOR
ENGINEERING
CORRECTIVE ACTION REPORT
QUALITY CONTROL PUNCH CARD SYSTEM USED IN MATERIAL REVIEW
QUALITY CONTROL REPORT
PUNCH CARD
QUALITY CONTROL CHART
PUNCH CARD REPORT
WEEKLY REPORT
SORTING AND TABULATING

AUTOMOTIVE DIVISION FIRST REPORT ON MACHINE TOOL CAPABILITY STUDIES

B. C. Jacob, Jr.
Minneapolis-Moline Company

Reduce the risk of wrong decisions and you implement progress. Such is the objective of the Task Force on Machine Tool Capabilities, as set up in the Automotive Division of the American Society for Quality Control.

Over the years, many machines have been a bone of contention between vendor and purchaser. When defective material was produced, the purchaser was prone to blame the machine; and the vendor to believe his machine capable, because he had seen successful sample runs. Such disagreements, even on a friendly basis, tend to increase the obstacles in the way of corrections and ultimate satisfaction.

Decisions on the part of a vendor or purchaser to spend correction money should be guided by proper facts. These facts must be reproducible and also must be accepted by the affected parties as reflecting a true picture.

As most of you have already discovered, existence of defects neither accuses nor absolves the machine. By the same token, a sample run without defects does not assure that the machine is satisfactory. Deviating product may be caused by inherent characteristics of the machine, by external factors, or by both. Unless a process produces essentially all deviating product, samples free of deviation are possible, and increasingly probable as the proportion of deviations decreases.

The Task Force proposes economical, reliable and reproducible methods which are to determine the variability inherently associated with the machine tool. Comparison of this variability with that allowed by specification, will show the amount of leeway for variation external (Chart II) to the machine. When this leeway is adequate, it is possible to run production essentially without defects. Excessive leeway is usually uneconomical. Inadequate leeway is usually both troublesome and uneconomical.

Establishment of a sound method to reach our objective required an understanding of several viewpoints. To achieve this understanding, the Task Force has included the Chief Engineer of two respected machine tool companies, the Master Mechanic of an automobile manufacturer, a Tool Engineer from another, a Statistician, a Chief Industrial Engineer, and four Quality Control engineers from several automobile companies. In addition, several guests were invited to contribute to our enlightenment.

The pooled experience of these professionals in different aspects of the field yielded many interesting and occasionally heated discussions. Divergent views were explored and usually reconciled. Each of us had beside the Task Force objective, an important objective related to his own viewpoint. Our efforts were to make those compromises which provide a good answer for all of us, with minimum disruption to any of us.

For example, no tool vendor would want to disappoint his customer by needless delay in delivery. Nor would he like to over-build his precision to the point he could not compete for business. He would like to be sure that unreasonable demands for service would not gain support.

The purchaser would not care to bind himself to acceptance through some scheme he did not trust. Production people would hesitate to delay delivery for some doubtful quality assurances.

Master Mechanics could challenge the quality capability when it failed to agree with their past experience. Conflicts might seem inevitable.

A statistician might shy from the approximations in a scheme which soundly meets the Task Force objectives.

From these varied backgrounds has emerged a fuller realization that these well guided compromises yield a valuable tool to enhance the quality-cost performance decisions. These decisions can be made with a much higher batting average through the use of simple economical means which have met the general approval of the committee.

To provide customers with acceptable quality at lowest cost, it is necessary to have proper balance in design, tooling, and process controls. Proper balance is unlikely if the contribution of each factor is not known.

Good design will carry specifications which can be met at a cost commensurate with the intended use. Tooling should be aimed at economically meeting such specifications. Operating controls must realize good usage of basic process capabilities.

The foundation for meeting these three requirements is knowledge of the basic process capability. Most uneconomical specifications come from lack of this knowledge. Failures to meet reasonable specifications result from (1) inadequate process capabilities and/or (2) inadequate operating controls. Again, only knowledge of the existing process capability can diagnose the cause or causes of the failure.

Inadequate tooling frequently comes from lack of knowledge of its capability at the time of its selection. It also frequently comes from failure to realize that operation cannot be perfect, but also adds to the variation (Chart II) even under the best of control. Stock improperly prepared for the operation, and machine wear are other causes of inadequate performance.

Poor control is often blamed on the machine. Only knowledge of the machine capability can show the true cause.

In short, knowledge of machine capabilities permits engineers to select specification tolerances enough wider than basic process spread to allow for reasonable control variation and a small amount of wear so that economical maintenance procedures may be followed.

Knowledge of machine capabilities and control allowance requirements avoids the dissatisfactions which result from inability to meet specifications, and the excess costs involved in over-tooling.

Realizing the important role played by knowledge of process capability, and the additional allowances needed for good control and economical maintenance, the Task Force has endeavored to investigate the reliability of several methods used to determine capability, and the problems involved in understanding and applying them.

Many of those who have tried to apply perfectly sound plans have noticed how many people seek all possible ways of casting doubt on any interpretation of data used in stating a machine or process capability.

At this point, it seems proper to emphasize that the Task Force is not trying to cause society at large to decide something that they do not now decide. Rather it is trying to provide that segment of society which already decides these things, with optimum procedures to reduce the risks inherent in such decisions. It recognizes that absolute certainty appears unattainable, and therefore proposes to provide economically feasible techniques to raise our certainty of good decisions to levels more acceptable than now typically exist.

All of the plans discussed are Quality Control techniques based on laws of chance. All have successful histories of practical use by many companies. Properly applied and interpreted all will yield improved results in most cases.

In each method, the production performance is predicted from samples taken under controlled conditions. The laws governing the relation between samples and lots are well founded mathematically. The way they operate falls well within the experience of most people in the general mechanical production field.

Armed with sufficiently sensitive gages, any illusion of identical sized pieces coming from a process is dispelled. We usually find a central tendency, or similarity of sizes, with some scattering larger and smaller. Measurements taken on large lots produced under relatively constant conditions usually have about 2/3 of the pieces in the central third of the total spread. Small samples seldom contain pieces near both limits of the lot spread. The average of sizes in individual samples usually resembles the average of sizes in the lot, more closely in large samples than in small ones. Likewise, the spread observed in large samples more closely resenbles the spread of the lot than is the case in smaller samples.

Statistical mathematics have provided us with simple ways of relating the conditions observed in samples to the facts of the parent lot, or production. They also give us measures of the reliability with which we can describe the parent lot from the information in our sample.

The reliability of our interpretations slowly increases as sample data increases. Data on averages become more reliable as the square root of the sample size. Spread data (by methods which look at all the sampled pieces) become more reliable as the square root of twice the sample size. Where we judge spread only by the spread between the largest and smallest pieces in our sample, we do not make much gain after the first few pieces in sample size, and actually are less efficient in large sized samples.

The last paragraph briefly states some statistical facts. The logic of these facts is apparent from many commonly accepted statements such as "first impressions often fool you", "don't put all your eggs in one basket", "experience is the best teacher", etc. The first says, in effect, further sampling will be more reliable. The second infers that losses in one case will be offset by gains in another. The third urges to learn more reliably by being influenced by more events, i.e. larger samples. Gambling odds and life insurance premiums are based on extremely large samples which are so reliable, that going broke in either business where the capital is large enough to survive the adverse sample is practically unknown.

The average, or high turnover stock in shoe sizes is well known by the experienced shoe merchant, as is the usual range of sizes desirable to stock. Judging by your own shoe size and the shoe sizes of those few friends whose show size you know, however, you would be unreliably equipped to stock a shoe store properly. Again, for reliability, sample sizes must be adequate for the intended use.

The statement concerning the efficiency of spread determination by size difference between the largest and smallest items in a sample refers to the comparative ability of ranges and standard deviation calculations in discerning variability in a lot from the data of a sample. Standard deviation calculations consider all the pieces in the sample, whereas range ignores all but the largest and smallest. Obviously in a two piece sample they are equally efficient. As sample size increases, we obviously ignore more information in using only the range, and we may rightfully suspect those ignored pieces of the sample are capable of adding to our knowledge of the lot if we know how and are willing to consider them.

The distribution of sizes in large random samples closely resembles the distribution in the parent lot.

To judge machine performance from samples our goal should be adequate reliability at least cost. Obviously, lack of pieces or man hours may force us to a lower reliability. However, failure to use the information from any piece measured is wasteful of measurement effort. For this reason, use of range in any but small samples is NOT recommended as a guide to capability. For larger samples either use of the average of the ranges in small sub-samples, or the use of probability paper may be recommended. Both methods are simple to use. Each has special advantages.

Use of the average range in small sub-samples where the pieces in each sub-sample are taken consecutively, and without process disturbance, gives high reliability in determining the basic machine capability for most operations, provided a large enough total sample is used. Where larger total samples are involved, the use of histograms, or better yet, probability paper adds an excellent picture of the distribution of sizes.

The Task Force recommends the use of the Shewart Control Chart for Averages and Ranges (also known as $\bar{X}$ and R charts) as the most broadly applicable basis for determination of machine capability. The total sample size normally should be not less than fifty (50) pieces, with samples up to two-hundred (200) pieces, being desirable where practically and economically feasible. The samples should be grouped in five (5)

piece sub-groups preferably run consecutively, but in any event run without process shift or disturbance in any sub-group.

To avoid waste, a capability run should be terminated if the difference in dimension between any two of the first five pieces exceeds 82% of the tolerance allotted to the machine or process, as this is virtual proof of excess variability. This is true since a range 82% as large as the tolerance predicts a process which uses at least all the tolerance, allowing nothing (or less) for any external factors. In fact a range over 61% of tolerance would be expected to fail the requirement of permitting 25% for external influences.

When a process qualifies under the above criterion, the ten (10) or more sub-samples are to be appraised against control limits. Out of control points usually should not contribute to the data used in appraising the process. If any substantial number of the points are out of statistical control, the following decisions appear proper. If the out-of control points are average points, the causes are satisfactorily explained, the dimension is operator, rather than process controllable, and the range data is in control, the process may be safely judged from the test. Only if the range data is generally in control, is the test safely interpretable. If the average data applies to process inbuilt functions, it also must be generally in control to be safely usable.

The useful data is to be interpreted by the standard method, where inherent process variability or spread is equal to six times the average of the sub-sample ranges, divided by d_2, which is 2.326 for five piece samples, or 2.579 times the average range. In most cases the process inherently will be capable of holding a set average plus or minus half the spread. This is typically satisfactory judgment since most observations in machine processing are reasonably similar to the mathematical normal distribution. A few types of observation such as off-square, out-of-round, etc. are skewed and may be decentered, so that they might be average value minus a third and plus two-thirds the spread, or some other such value.

If the total data is in control, or most of it is, and those point out of control are discarded, a histogram, or better yet, a plot of the data on probability paper will reveal whether the assumption of symmetry and/or normality is safe.

When the process or machine inherent spread has been determined, comparison of this spread to the part print tolerance may be made. A generally accepted view holds that the inherent spread should not exceed 75% of the part print tolerance. In especially critical cases it might be desirable to allow it less than 50% of the specified tolerance. In no case may it equal or exceed the specified spread, since tool wear, tool setting and other such unavoidable disturbances will always increase the operating spread beyond the inherent spread. (Chart II).

Where a dimension is inherently machine or process fixed, as a distance between two holes simultaneously bored by a pair of fixed spindles, the average determined by the test, plus and minus half the spread calculated from the range data must fall within specified location tolerance, preferably with some margin for maintenance.

These recommendations reflect careful consideration for simplicity and dependability, factors typically in conflict. The use of ranges

instead of the longer more rigorous mathematics only slightly increases the possible error in an estimate of the process performance, but greatly simplifies obtaining a reliable estimate.

When we speak of reliability, we mean assurance that our opinion of capability is sufficiently exact, in the vast majority of cases. This may sound like hedging. We want to be "exactly exact" and "every time". Unfortunately we know of no human being who knows a means of fulfilling this want. We can, however, compare the reliability of our proposal under some varying conditions. For instance, most test values will have been calculated from a sample which had more than typical variation or less than typical variation.

In small samples, extreme dimensional values of individual pieces will distort our opinion of the lot or production more than such values would if counter balanced by a larger number of typical pieces. The sample calculated value of an individual sample might occasionally give a very distorted picture of the expected performance. Yet if we are willing to be right 90% of the time, we can state that for a given size of sample, the lot will have variation within $\pm$ or — fixed percentages of the sample indicated variation. If we want to be right 98% of the time, the $\pm$ or — values are larger, but equally definite. For 100% of the time no one can provide a definite limiting error in opinion.

The band of predictable difference from sample derived variation widens very rapidly as total sample goes below 25. It narrows quite slowly as total sample size increases over 50 pieces. What does this mean to us? First, it means that samples less than 25 should be used only when no other alternative seems practicable, and then one must realize that his opinions contain much doubt. Second, the requirement of samples over 50 pieces should only apply where larger samples involve reasonable costs.

Chart I shows the limits of variation from sample estimated spread to be expected in a lot with sample sizes from 2 to 100 pieces when using range in sub-samples as a basis. Although the choice of ranges differs slightly from those recommended by the Task Force, the values plotted are similar enough to justify use of this chart for guidance purposes.

Only slightly increased reliability is available from any other system, and that is achieved with far greater complexity and effort. Assurance that the sample properly represents the lot is better than with most other methods. Methods which are even slightly simpler than the proposal suffer substantial loss of reliability.

The objective of the Task Force, namely, reduction in the risk of wrong decisions, through economical, reproducible and dependable methods of appraising the variability inherently associated with the machine tool, appears to be met by the proposed system.

The system is summarized as follows:

1. A total sample of 50 to 200 pieces depending on economic factor is recommended. These should be so run that no basic disturbance enters the run of any single 5 piece sub-sample.

2. When the dimensional difference between any 2 of the first 5 pieces exceeds 82% of the specified tolerance the test should be stopped as this is virtual proof of excess variability.

3. If condition 2 is met, plot for each 5 piece sub-sample the range, which is the dimensional difference between the largest and smallest observation in that sub-sample.

4. Plot for each 5 piece sub-sample the average ($\bar{X}$) of the sizes in the sub-sample. This is simply computed by adding the 5 dimensions, and multiplying the sum by 0.2 (i.e. double and move the decimal one place left).

5. Compute the average range ($\bar{R}$) by adding the sub-sample ranges (R) and dividing the sum by the number of sub-sample ranges used.

6. Multiply $\bar{R}$ from 5 by 2.11 for an upper control limit on ranges, and compare each individual sub-sample R to this limit. If all are below this limit the range data R are all acceptable. If any are over the limit, those sub-samples should not be used. If the over control ranges are a very minor proportion of the observed sub-samples, or if the cause of disturbance of the over control ranges is known and known to be external to the machine, proceed as follows:

   (a) Recompute $\bar{R}_7$ (meaning the $\bar{R}$ for use in step 7) as in step 5 except do not include any data from over-control sub-samples.

   (b) Repeat step 6 using the revised $\bar{R}_7$ of 6 (a).

7. Using the correct $\bar{R}_7$ (from step 5 or 6 (a) as applicable) multiply by 0.58 for A_2. Compute grand average dimension $\bar{\bar{X}}$ which is the sum of the $\bar{X}$'s of step 4 divided by the number of sub-samples observed. Check to see if each $\bar{X}$ lies between $\bar{\bar{X}} + A_2\bar{R}_7$ or $\bar{\bar{X}} - A_2\bar{R}_7$. If all do, all average data are acceptable. If the out of control averages are a very minor proportion of the observed samples, or if the cause of the disturbance is known, and known to be external to the machine, or operator adjustable in the machine, the data are acceptable. If the basic dimension is not operator adjustable, or a distinct exact external cause is not known, no safe conclusion may be drawn.

8. When the above steps qualify the sample as representative of the performance of the machine, the following capability determination is made.

   (a) For operator adjustable dimensions, the $\bar{R}_7$ multiplied by 2.579 is full inherent process spread and may be compared to specified tolerance spread to allow for variations outside the machine. Process knowledge would indicate that in some cases, it should be substantially less than 75% of specified tolerance spread, as in certain grinding operations where wheel wear would otherwise require excessive attention to setting.

(b) For machines which are not operator adjustable for the observed dimension, the machine capability is judged to have the limits $\bar{\bar{X}}$ of step 7 $+$ or $-$ 1.29 $\bar{R}_7$. These limits may be compared to specified limits. It is recommended that some margin be allowed for error and external factors. Where 2-sided limits apply a similar basis to step 8 (b) would show that $\bar{\bar{X}}_7$ $+$ or $-$ 1.6 $\bar{R}_7$ should fall inside the specified tolerance. Where the tolerance is one sided, such as only a maximum or only a minimum specified, it will be safer to compare $\bar{\bar{X}}_7$ $+$ or $-$ 1.29 $\bar{R}_7$ (as the case may require) with 75% of the specified maximum or 133% of a minimum.

References

1. Grant E. L., Statistical Quality Control, McGraw-Hill, New York, 1951.

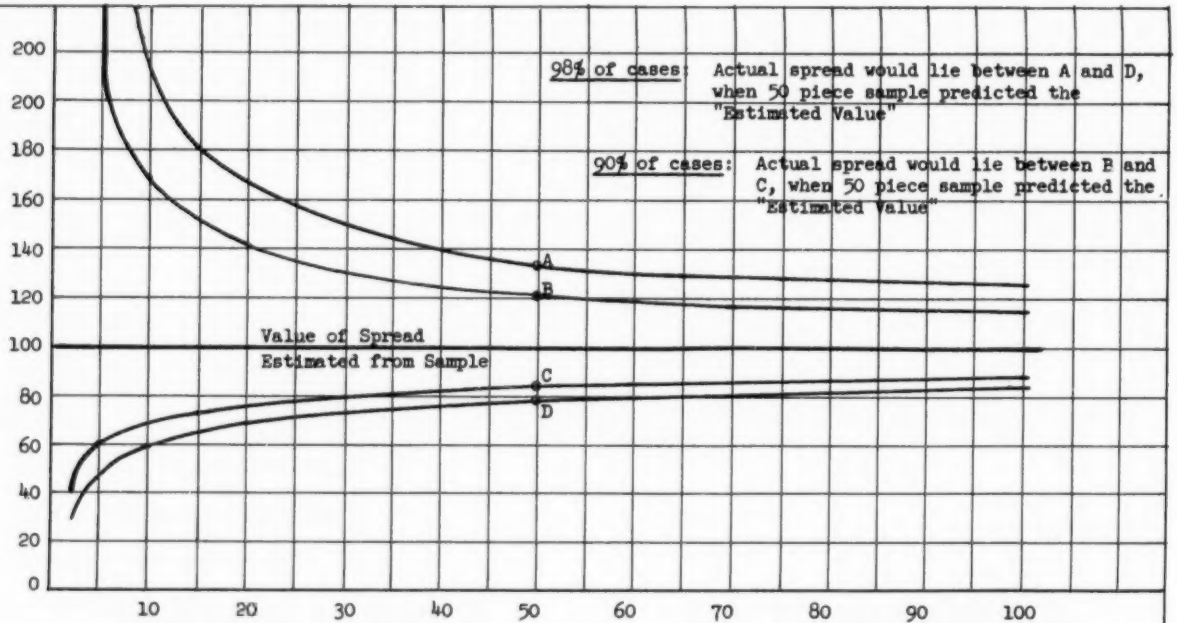
2. Bayer, Harmon ("Quality Capability of Machine & Tools") (Tooling and Production, May and October, 1954).

3. DeGrote, Irwin A."Try it for Size" (Measuring Machine Performance) Paper presented to The National Machine Tool Builder's Association, Houston, Texas, March 29, 1956.

4. Frank E. Grubbs & Chalmers L. Weaver "The Best Unbiased Estimate of Population Standard Deviations Based on Group Ranges" American Statistical Association.

5. Arthur Bender "A Quick Method of Determining the Characteristics of Frequency Distribution" 8th Midwest Quality Control Conference Papers, 1953.

6. Gottfield E. Noether, article American Statistical Association Journal, December 1955.



Lines A, B, C, and D plotted from data in Table I, American Statistical Association Journal, December 1955, article by Gottfried E. Noether, page 1046.

CHART SHOWING LIMITING PREDICTED LOT SPREADS AS A PERCENT OF SPREAD ESTIMATED FROM A SAMPLE

Chart I

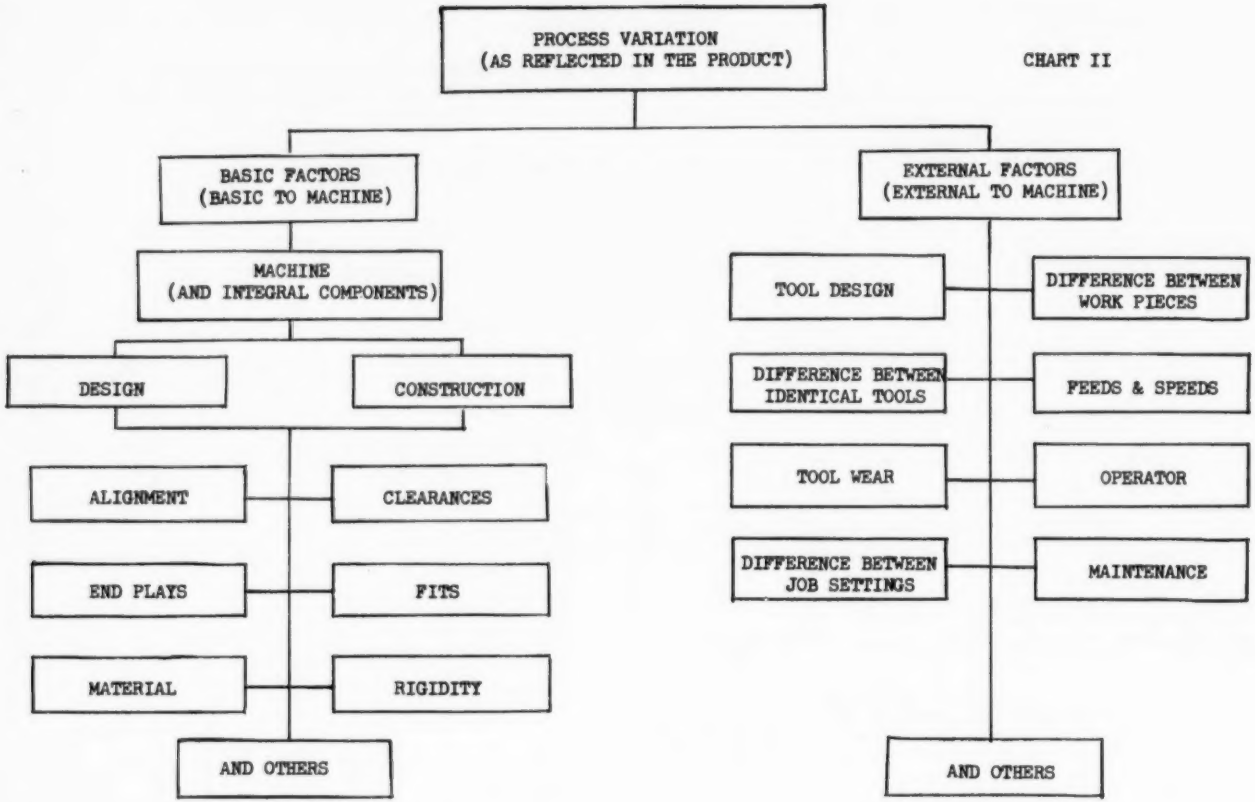

PROCESS VARIATION
(AS REFLECTED IN THE PRODUCT)
BASIC FACTORS
(BASIC TO MACHINE)
MACHINE
(AND INTEGRAL COMPONENTS)
DESIGN
CONSTRUCTION
ALIGNMENT
CLEARANCES
END PLAYS
FITS
MATERIAL
RIGIDITY
AND OTHERS
EXTERNAL FACTORS
(EXTERNAL TO MACHINE)
TOOL DESIGN
DIFFERENCE BETWEEN WORK PIECES
DIFFERENCE BETWEEN IDENTICAL TOOLS
FEEDS & SPEEDS
TOOL WEAR
OPERATOR
DIFFERENCE BETWEEN JOB SETTINGS
MAINTENANCE
AND OTHERS

CHART II

QUALITY PROBLEMS OF AUTOMOTIVE FABRICS FROM THE PRODUCER'S POINT OF VIEW

Everett C. Atwell
Burlington Industries, Inc.

No manufacturer can expect to survive in today's intensely competitive market without a healthy respect for a strong quality control program. The maintenance of full control of raw materials entering a process and at each step during processing is necessary to assure compliance with end use specifications and to keep at a low level, seconds, unmerchantable goods, rehandling and lost productive time.

The first step in insuring satisfactory quality in a given automotive fabric must start with the fabric construction itself. A combination of fibers and yarns must be chosen to provide aesthetic appeal and a minimum of production quality problems. It is here that the story of quality control begins. In the initial design of a fabric, a base fiber is chosen which will provide the strength and stretch to withstand repeated loading without early fatigue. A second, or a second and third fiber is selected usually to form the face of the fabric which combines strength, stretch, and abrasion resistance and with dyeing properties which lend themselves to unique color combinations with the other fiber. The fibers that are selected must be characterized by good spinning performance in the case of staple fiber and provide satisfactory yarn appearance and dye levelness in the finished fabric. Quality control is exercized all the way from checking the uniformity of the staple fiber stock going into work to the finished yarn. These tests include checks on weight, uniformity of weight, twist and breaking strength. Filament yarns received by yarn producers are, by experience, so well controlled for quality in production that a weaver seldom finds it necessary to conduct more than random tests on incoming shipments. Weaving operations must be fully controlled and every precaution taken to prevent visual and latent defects. This is especially true in preparatory work on filament nylon where uneven tensions or other process irregularities can cause unsightly streaks. Finally, after the first production of greige fabric is woven, each roll is given 100% inspection and graded as standard or substandard. At this point, all conditions which have been major contributors to the production of substandard fabric which cannot be shipped to the customer are immediately followed back with the appropriate production section for follow-up and correction. In many instances the cause or origin of the type of defect observed is obvious while in others a fairly comprehensive group of tests must be made in the Laboratory to ascertain the source of the trouble - following which remedial steps can be taken.

Automotive production units, as in a majority of other types of manufacturing, set up once each year to produce a limited number of kinds of items. Production usually runs on these items for a full year without more than a few corrective changes in design or standards. In contrast, the successful textile production unit today is faced with setting up for short runs on a multiplicity of styles which puts pressure on the quality control group to detect defects and apply corrective action with the utmost speed, otherwise, it is possible to get into a situation of giving late deliveries and building up an inventory of substandard fabrics with no outlet other than at prices under cost. Thus you can see that the observation and fast correction of defects resulting in inferior quality

is the key to a high standard of quality not only in greige goods but also in finished goods deliveries to customers. In fact, the rapidity with which these corrective steps are taken is frequently the key to whether a supplier can do a satisfactory production job or not.

All large manufacturers have at least two sources of supply for each item purchased. This means that for each new style accepted by an automotive customer at least one other supplier must also satisfactorily and quickly duplicate the style. This involves analysis of the competitors sample, design, the preparation of greige goods, finishing of the sample piece and submission to the customer. This complete waste of time, added expense and effort on the part of the suppliers has been recently resolved by the free interchange of construction details between successful bidders to the automotive trade.

Whereas in the production of automotive greige goods, greater emphasis is placed on inspection techniques and mechanical controls, greater emphasis in the dyeing and finishing operation is given to chemical controls. These consist of checks on detergents used in the boil off, dyes, wetting agents and dyestuff fixatives used in the dye bath, and functional finishes such as water repellents, resinous face finishes and back sizing used in the finishing operation. Additionally, the application of the materials selected for use are subject to control. This includes time, temperature and concentration of detergents in the boil off tank, time, temperature and quantities of dye and auxilliaries in the dye bath, total solids in the back sizing and temperatures and width in drying.

The final quality of the fabric prior to shipment to automotive accounts is determined first by 100% visual inspection and classification and secondly by adequate sampling of the rolls for subjection to end use tests. It is not until inspection, classification and end use tests have been completed that a lot can be considered for shipment.

I think we would all agree that the major aim of suppliers to the automotive trade is to deliver goods that meet the purchasers requirements. One problem which confronts us is the dissimilarity of performance requirements in these fabrics, all of which are going into the same basic end use and hence should have the same requirements. In the field of fabric inspection, which is one phase of the quality control picture, there is no particular problem inasmuch as here we are concerned with defects which are readily apparent to the eye and, while our judgment may vary to some degree, requirements are essentially similar so that there is no great difference of opinion. However, in the case of end use testing, the automotive specifications which the suppliers are required to meet differ not only in terms of the test method but frequently where the test method is similar, the evaluation is so different as to make the level of acceptability quite different. For an example, I might mention the matter of testing a fabric for resistance to sunlight exposure. In this particular case the supplier to Company "A" is required to meet a 90 hour exposure to Florida sunlight while Company "B" requires that the fabric pass a 50 hour exposure in the Fadeometer and Company "C" requires the fabric to pass a 90 hour exposure in a Weatherometer. The results of testing a fabric by one procedure differs from the results obtained by the others. This simply means that a fabric that is acceptable to one customer for a specific end use may be wholly unacceptable to one or both of the second and third customers for the same end use. Then there is the situation where there is no acceptable test method available to

evaluate a fabric for some particular service or function which it must perform or to meet a condition to which it must be resistant.

Here again the supplier is obliged to test deliveries to each customer by the customers' methods which are usually quite different and again the same goods may be acceptable by one method and unacceptable by the others. While there is no intention of being unduly critical of the individual tests themselves, we do feel that unanimity of such procedures would be to the best interest of all concerned. Suppliers would then have a single condition to meet which would put them on much firmer ground when fabrics are made on a sample basis and it is not known at that time to which customer the submission is to be made.

I have no desire to labor the point but the same discrepancy in requirements exists in many other areasin the specifications and we feel that a great deal could be accomplished in achieving better quality deliveries through standardization of testing procedures and evaluation on automotive fabrics. The very obvious thought will probably occur to many as to why do not the suppliers shoot for the very maximum in end use performance and thereby avoid any field of controversy by establishing our own standards so high as to encompass all. The answer is equally obvious to those of us who have had experience in trying to meet the requirements of this trade. It is simply the age old problem of not being able to improve one characteristic without forcing some other feature to suffer thereby. As light fastness is improved to a desirable level quite frequently trouble will be encountered with leaching or a transfer of color within the fabric on wetting. When proper chemicals are used to correct this deficiency, the light fastness may be affected considerably and so it goes.

The quality problems of the sort mentioned above are fairly common to all the automotive suppliers. In certain instances requirements which we may feel are unduly high and by the same token extremely difficult to meet, could make for considerably better overall performance if the level required in the aforementioned instances were reduced. In an effort to attempt to focus attention on certain constructive action which should improve the overall quality of goods supplied by the various processing firms to the automotive trade, an Automotive Upholstery Fabrics Association has been formed. This association consists of representatives from each of the major automotive fabric suppliers and its prime purpose is to consult on mutual problems, come to agreement on a course of action which would be mutually acceptable and practical and present both the problem and recommendation to the various automotive concerns for their consideration. We believe that this will be a very constructive step if these recommendations are given serious consideration and free discussion.

Another approach to the problem of developing test procedures to determine characteristics or performance under conditions for which standard procedures are not now existant involves consideration of presenting such problems to either American Society for Testing Materials or the American Association of Textile Chemists and Colorists. These organizations are comprised of technologists from all segments of the textile industry. Task committees are formed for each special problem comprising specialists in the fields covered by the problem at hand. These committees formulate most of the standard procedures used to characterize behavior of textile materials under various conditions of use. Procedures formulated by the AATCC do not become standard procedures for

industry use until the procedure has been in use in a "tentative" status for at least one year and then only by a majority vote first by the Technical Committee on Research and then by the Executive Committee on Research. Essentially similar safeguards are practiced by the ASTM. It is to be especially noted, however, that these are strictly test methods as it is not a function of either of these organizations to suggest, promote or establish minimum or acceptable quality levels. The adoption of this approach to the formulation of new test methods would provide a unity of purpose and procedures acceptable to all concerned without removing the prerogative of the automotive companies to establish their individual levels of quality acceptance.

Where large numbers of like units are being produced, statistical methods of control can be employed with considerable savings, but even more important, better assurance of uniform quality. Where these conditions are not met and we have multiple small quantities of many dissimilar units in production, statistical methods are impractical if not impossible to apply in our experience to date. Statistical methods are more readily adaptable to the control of laps, roving and finished yarn production but are increasingly difficult to apply at points beyond the finished yarn stage. For example, application of the method can be applied to tension controls and to establishing causes for yarn breaks in the preparatory area and to the evaluation of major and minor causes for visual defects. In some instances it can be applied to fabric inspection, however, this has not as yet been adapted to automotive fabrics primarily because we are not only interested in determining average level of quality for control purposes but it is of major importance that individual pieces be properly classified to prevent shipment of substandard goods to our customers.

In summation, it should by now be apparent that the quality control of production of automotive fabrics is a very challenging field by reason of the high requirements of quality and performance which are much more difficult to attain than for the majority of the end use requirements in other fields. Great strides have been made over the past ten years by way of more appealing styling, greater functional performance and greater durability. We, the textile suppliers to the automotive industry look forward with optimism to our ability to keep abreast of both technological advances and changes in customer requirements in maintaining a high standard of quality in this field.

QUALITY CONTROL TRAINING COURSES
PART ONE
IN-PLANT COURSES

Joseph Movshin
Thompson Products, Incorporated
and
Rudolph Freedman
Alco Valve Company

In discussing In-Plant Training Courses in statistical quality control, we must consider the background of the individual we are going to train, his place in the organization, the use to which he will put statistical quality control, and his interests. These factors may be viewed in connection with three of the catagories of people within an industrial organization.

A. Production Supervision
B. Practicing Quality Control Technicians
C. Design Engineers

When we refer to production supervision, we usually have in mind a "foreman". This man is usually charged with getting production out and in order to do so, it often becomes necessary for him to train workers and to work with other staff departments in the application of many of the industrial techniques that now associate themselves with production. He has no natural interest in statistical quality control and to him it is just another industrial tool. Indeed, his interest must be aroused in order to take statistical quality control out of the fad stage and into the truly useful tool stage.

Before the method of training is discussed, we must decide the purpose of that training. Do we want to make a Quality Control Engineer or do we want to teach the man enough about quality control that he will merely be acquainted with it? In view of the duties that a Production Supervisor already has, I do not think that we want to make a full fledged Quality Control Engineer. Also, his background does not readily lend itself to training him fully in the techniques of statistical quality control. However, I think that we want more than merely to acquaint him with the subject. It is important that a Production Supervisor knows enough about quality control that he will know its limitations as well as what can be accomplished by its use. He should learn enough to work intelligently with the Quality Control Department.

The actual training should start with a sales program to arouse interest in statistical quality control. After interest is aroused, the sales program must continue until the supervisor is convinced of the usefulness of the technique. This means that the entire training is, to a large extent, a selling plan. A good start should take the form of individual talks between the Quality Control Engineer and the supervisor. Reference might be made to how statistical quality control worked in other companies and how these other companies have profitted from the program. It should then be pointed out that the program can help not only the company but also the supervisor's department and the supervisor himself. A display of interest-arousing gadgets, that demonstrate statistical quality control principles, is helpful. Allowing the supervisor to use these gadgets and to show them to some of the people who

work for him will often make him want to learn more. There are a number of such items available.

After his interest is aroused, he should be given a course in basic statistical quality control. In order that there might be adequate group discussion, not less than four should take the course at one time. If the instructor is to give each the proper attention, there should not be more than fifteen in the group.

There are certain features that the course should have.

1. It must be simple. It must be given in terms that are easily understood. The mathematics should be kept at the absolute minimum. The use of shortcuts such as tables and calculators should be encouraged.

2. It must be basic. Only such things as frequency distributions, control charts, and acceptance sampling by attributes should be covered. No special techniques should be introduced. If there is a need for these techniques, they may be introduced at a later time.

3. It should be built around the principle that they learn by doing. Emphasis should be placed on allowing everyone in the course to take an active part. This may be accomplished by the use of work sheets, which should be completed in the classroom. Whenever possible, the data for these work sheets should be obtained by measuring a characteristic on actual parts being used or manufactured by the company. One of the characteristics of these parts may be measured by the group itself. A frequency distribution, based upon these measurements, usually will bring about a full amount of discussion about the nature and causes of variation. Work sheets should be of the type where proper blank spaces are provided for the data and the results. The instructor should make sure that each one is working the problem correctly. The results obtained should then be analized with the instructor acting more as a discussion leader than as a lecturer. It should be shown how the principles learned in connection with the work sheet apply in industry.

4. There should be very little or no homework. Most of the work should be done in class. If homework assignments are given, and the supervisor gets stumped at how to work out the problem, he is liable to spend hours arriving at a solution, and might very well end up feeling frustrated. He will then build up a defense against statistical quality control. Therefore, homework assignments should be very simple. Far better than a homework assignment, is for the Quality Control Engineer and the supervisor to obtain data and analize the data together. The results can then be presented to the class for discussion.

Such a course will take about six or eight sessions of about one and one-half hours each. However, it is important that the instructor does not have the feeling or attitude that certain subjects must be

covered within a certain time. It is far better to compromise with the amount of material covered than to have some members of the class fail to get full understanding of that which is being taught to them.

In the case of Quality Control Technicians, we are referring to those people who carry out the details and the leg work of the programs that are set up by Quality Control Engineers. Some of their typical duties might be to collect data, tabulate that data, make control charts, advise others of routine findings from control charts, select a sampling plan for specified conditions, etc. Very frequently they are members of a union bargaining unit and they might have held various production jobs in the past.

The training of these people, at first thought, should really be quite simple, as it should be a case of self-interest. In other words, we might assume that indoctrination would not be necessary as the man who is going to use quality control would realize that it is in his self-interest to be trained in the methods. Generally, this is true. However, all too frequently we encounter individuals who will gripe and grumble about anything new that is taught to them even though it would serve their best interests to learn and apply these principles. Somehow or other, these individuals must be reached before they are turned loose in the Production departments or their griping and grumbling will erase any of the good work that is being done by the other Quality Control Technicians. They must be convinced that it is in their best interests to learn and do a good job of application.

If we assume that we have already convinced all of the people involved or that we have had in the first place a receptive group, we will still have a great problem with diverse backgrounds. Some will have the education, the background, and the experience that will enable them to learn very quickly and easily. Whereas, others will be very slow in grasping what is being taught. Therefore, we have the age old problem that confronts every educator, namely, to go fast enough in the course to keep the better ones interested but still to go slowly enough so as not to lose a number of the students. Fortunately, in any in-plant course, we have the advantage of being able to stay in contact with the students all during the working week. Therefore, tutoring of the slower people is not at all difficult or out of order.

The course itself will depend upon the need and the application that the company has in mind. In some cases, simple applications of frequency distributions, control charts, and acceptance sampling by attributes is sufficient. Although the course is elementary, it is important that it be thorough. Sympathetic understanding is sufficient in the case of production supervision but not so with quality control technicians. They must be taught in great detail as this is the group that will form the link with other departments. In some cases, it will be desirable to teach special techniques and go into the background information that goes with the fuller knowledge of quality control.

It is desirable within this group to encourage attendance at A.S.Q.C. local section meetings, and Regional Conferences. The most outstanding people in this group, depending upon their position within the company and their grasp of the subject, might even be sent to courses that are offered by some of the A.S.Q.C. sections and by some of the universities.

When we consider the In-Plant Training of Engineers, we are thinking particularly of Product Design Engineers and Process Design Engineers. The educational background of these people is usually adequate for easy learning of the basic elements of statistical quality control. Even many of the special techniques are absorbed without too much difficulty. However, there are often several forces operating to limit interest in their application of statistical quality control. Very often they are under pressure to make a prototype and then get production going. They are not allowed enough time to gather data for even the smallest of samples to which statistical analysis may be applied.

Some large companies have set up quality control sections and programs within the engineering departments and this serves to free the other engineers from testing programs as soon as possible. This section may also be used to set up methods for production quality control.

The training program for engineers may be very informal. A few group discussions will help to stimulate interest. Information on the actual techniques may be transmitted by making available a few well chosen articles and books. A few joint applications of the administrator of the training program and the engineer will serve to get the latter started.

QUALITY CONTROL TRAINING COURSES
PART TWO
ASQC SECTION COURSES

Joseph Movshin
Thompson Products, Inc.
and
Rudolph Freedman
Alco Valve Company

Since the very beginning of the ASQC, local sections and specialized groups of the Society have conducted training courses in Statistical Quality Control. In the 1955-56 season, some forty sections of the Society offered sixty-two different courses. While exact figures are not available at this moment, it is estimated that at least fifty sections will have given about one hundred courses in the 1956-57 season. In general, these courses have three basic purposes:

1) To offer general introductory material on Statistical Quality Control. (Sources of such information are not generally available and in many instances ASQC courses are the only effective way of introducing Statistical Quality Control to local industry).

2) To provide a means of improving the technical competency of members of the Society and others in the field of Statistical Quality Control.

3) To provide a means of contacting people who might be prospective members of the Society. This is accomplished through the attending publicity as well as through participation in the courses themselves.

Often a course is designed to accomplish to some degree all three of the purposes stated.

As might be expected, the material presented has been somewhat dictated by the objectives of ASQC Training Courses. In general, they tend to be the "How you do it!" and "How good is it for you!" sort of thing without too much emphasis on the technical fundamentals or statistical methods involved. In many instances, "selling" is an important part of the program. Of course, there are certain specialized courses which attempt to provide sound fundamental statistical training. Generally, however, these are in specialized fields and are considered as "advanced" or "special" courses offered to select groups.

The underlying purposes behind the ASQC Section Courses have resulted in certain characteristic attendance. In general, the classes will be quite heterogeneous and cover a broad cross section of industrial and educational backgrounds. A single group will often have some operating shop personnel with perhaps highschool background. These will be interested only in acquiring some understanding of what SQC is all about. Inspection and quality control personnel of similar educational background will be interested in the mechanism of making and using control charts and sampling plans. There will be engineering graduates

who want to know something about this technique, and will also want to understand more about why it works. Administrative and executive personnel will be interested in what quality control can do for their Company, and how one goes about getting these benefits.

This, then, is a broad picture of the ASQC Section Training Course. The detail of course content will not be considered as such in this paper. We will merely discuss the organizational and administrative aspects of an ASQC Section Training Course. The material is presented in outline form to attempt to highlight the points to be considered. The information given is based on an analysis of courses offered by all the sections of the ASQC as reported to the Education and Training Committee from 1955 through the Spring of 1957. It includes ten years experience of the authors as organizers and teachers of courses with the St. Louis Section, ASQC, and Washington University in St. Louis. Administrative problems were also clearly pointed up to the authors; both of whom are past Presidents of the St. Louis Section, ASQC.

1) Course Selection

In the past, the course selection has often been just another matter for the local Education Committee or Executive Committee to decide. Long-range programming was generally not considered. It became habit to repeat successful courses on an annual or semi-annual basis and to discontinue courses that proved relatively unpopular. Recently, however, programming for a longer period of time has been adopted by several sections. Specifically, a series of courses are developed and outlined. These are then offered sequentially so that an individual can take a basic or introductory course and, can then take other courses to develop his skill or background. An example of such a series is that recently developed by the St. Louis Section. This includes the following five courses:

A) Introductory Statistical Quality Control
B) Intermediate Statistical Quality Control
C) Sampling
D) Pre-Advanced Statistical Quality Control
E) Special Techniques

The first of these is purely introductory. The second describes the basic control charts in slightly more detail with some of the statistical theory involved. The third discusses attribute and variable sampling plans and some of the probability theory underlying standard sampling tables. The fourth develops distribution theory and basic statistics to the point where it can be applied to certain specific techniques such as analysis of variance, correlation, or design of experiments. The Special Techniques course is more of a "How you do it!" sort of thing with some theory on what are sometimes called "advanced techniques". These include: analysis of variance, correlation, chi square, design of experiments, specialized control charts, and similar topics. The Special Techniques course can be altered from year to year depending on popular interest and demand. Each of these is presented in ten sessions of one and a half hours. At least two courses in the series are given each year -- one in the Spring and one in the fall. If the demand warrants, two may be given concurrently for a total of four a year.

A) Subject Matter

While the subject matter for ASQC Section Courses could theoretically vary widely, it has become somewhat standardized in the past few years. About 75% of the courses given are basic or introductory in nature. The remaining 25% are about equally divided between courses called "Intermediate or Specialized", and "Advanced". The so-called "Intermediate" and "Advanced" vary greatly; the words having meaning only with respect to the level of activity in the particular section and the general state of the art in the area.

Several sections have conducted surveys of their membership and of industry in their areas to determine what courses or subjects might be of interest. Information from such surveys can prove very helpful. Such surveys are recommended to all sections to periodically determine the changing interest of the community. The questionnaire should consider both academic SQC areas as well as so-called "How you do it!" subjects. Surprisingly enough, surveys the writer has seen tend to favor the academic rather than specific "How I make an individual control chart!" sort of thing.

B) Number of Hours

Courses offered by ASQC sections vary from 1/2 hour in length to 3 hours in length. The great majority of courses are 1 1/2 to 2 hours in length per session. It is the writers' experience that after 2 hours and possibly after 1 1/2 hours, the mind can no longer endure the "suffering of the seat". Once this point is reached, the mental absorption of external matter is reduced to the vanishing point.

The total number of hours per course varies greatly from under ten upwards to sixty. There is no predominant favorite although the modal value is between fifteen and twenty hours. Ten sessions of about 1 1/2 hours is a popular average for the introductory course.

C) When and Where Given

In general, there is no clear preference in this. About one-half of the courses are given at a college or other educational institution. The other half are given at a hotel or other public meeting place. This does not seem to be an important point and is probably a matter of local convenience. The same is true of the time. Equal success seems to have been met on late afternoons (4:30 or 5 o'clock), evenings, or Saturday morning courses. It might be noted that failures can also be reported at these times. It seems in general that the local interest that can be developed is more dependent upon course content, instructors, and publicity than upon the exact details of time and place. It is obvious, of course, that conflict with other activities should be avoided in so far as practical.

D) Text

The overwhelming majority of the courses given used

little or no text material. The following is a distribution of materials reported for 1955-56:

Papers & Manuals	No. Reported
1. Local Notes and Papers	16
2. "A Basic Training Manual on Statistical Quality Control" (St. Louis Section, ASQC)	11
3. Mil-Std. 105-A	8
4. "Quality Control Training Manual" (Iowa Section, ASQC)	3
5. Extension Course in SQC by Westman	2
6. General Motors Inspector's Training Manual	1
7. Nav. Ord. OSTD 80	1
8. ASTM Manual on Quality Control of Materials	1

Books	
1. "Statistical Quality Control" by E. L. Grant	13
2. "Statistical Analysis" by Dixon & Massey	3
3. "Facts from Figures" by J. M. Moroney	2
4. "Statistical Analysis for Chemistry & the Chemical Industry" - by Bennett & Franklin	1
5. "Industrial Experimentation" by K. A. Brownlee	1
6. "Engineering Statistics & Quality Control" by I. W. Burr	1
7. "Statistical Theory with Engineering Applications" by A. Hald	1
8. "Inspection & Gaging" by C. W. Kennedy	1
9. "Introduction to the Theory of Statistics" by Yule & Kendall	1

It is the writers' opinion that more formal text material should be given to the students at these courses. Even though this might not be used in the course, such material might prove an incentive to those taking the course to study further in the field. It also helps the student feel that the course is worthwhile and that there is a scientific background to statistical quality control even though it is simplified for daily shop use.

E) Tie-in With Special Groups

Often courses are given to appeal to a local special interest group. This might be an industrial or trade group. In these cases, the course is specifically adapted to the purpose at hand. In general, these are given by ASQC Sections on special request rather than part of a general education program.

2) Instructors

In general, it is felt that section courses should be on as professional a basis as possible. To this end, the instructor should be an authority in the field of SQC and possibly a professional educator. Since such a person is not always available, a well-known local practitioner of quality control or a professional educator in mathematics or statistics serves. There are generally found within the ranks of section members or local educational

institutions. Generally, more than one instructor is used even for a limited course; although the practice of using too many instructors tends to destroy course continuity. Many sections prefer two or three instructors for a 15 to 20 hour course to create interest and variety.

While general practice varies, it is felt that instructors should be paid a nominal amount to maintain section courses on a professional level. An average honorarium that has been used is approximately $20.00 for a hour and a half session. It is necessary, however, to note that many of our best section courses have been taught by the most competent men in the Society at no charge. The fact that such men are not generally available to many Sections, however, adds emphasis to the need for a rigid professional approach to the problem of instructors.

3) Budget

While the purpose of ASQC Section Courses is not primarily to make money, it is necessary to consider cost and budgets. The following is a typical budget that might be considered for a section course. This represents average costs of courses given in the St. Louis area for a ten session course of one and a half hours each:

Estimated Income

10 Registrants @ $20 (members of ASQC)	$200.00	
10 Registrants @ $30 (non-members of ASQC)	300.00	
		$500.00

Estimated Expenses

Hotel (10 nights @ $12.50)	$125.00	
Instruction fees	200.00	
Training Manuals 20 @ $1.00	20.00	
MIL-STDS 105A 20 @ 30¢	6.00	
NAV ORD Variable Sampling Plans 20 @ 50¢	10.00	
Printing	50.00	
Postage	20.00	
Miscellaneous	5.00	
ASQC for 10 memberships	60.00	
		$496.00
Net Favorable Balance		$ 4.00

Note that ASQC members are charged an average of $20.00; non-members $30.00. The extra $10.00 for non-members, however, often is used to secure a membership for these people.

4) Membership Tie-in

About 20% of the sections reported a membership tie-in with their training course charges. While a great many arrangements were reported, generally, non-members were charged a $20.00 fee with about a $10.00 additional charge for non-members. In almost every case, this included a membership in the ASQC although several sections reported an option on this item. There seems to be no clear cut

opinion as to the desirability of this practice. It is prevalent and has been successful enough in obtaining members for many sections to be worthy of consideration. In several instances, however, it was felt that members so obtained did not represent truly interested and permanent members and as such, were not desirable.

5) Brochure and Publicity

The brochure often has a dual purpose: to announce the course to the individual that receives it; and possibly to be used as a poster on bulletin boards, etc. It should be efficient and professional in nature, but not garish or gaudy. Generally, 8 1/2 x 11 paper which may be folded twice for mailing is used. The front may be printed from top to bottom and contain all the important information laid out as a billboard announcement. The back is devoted to details, registration form, and mailing section. Of course, other folders or book forms have been used successfully. It is generally desirable to work with local printing or advertising people who can advise and assist in this manner.

It is extremely difficult to obtain enough publicity. The brochure should be mailed at least to all the industrial firms of the area and to all members of the Society. In certain instances, brochures are mailed to an entire state or inter-state area. Frequently, Chambers of Commerce of communities within a radius of fifty miles or other industrial trade groups in the area will assist. Technical societies may participate and provide mailing lists or include announcements in their mailings. Local colleges or universities and educational groups will often assist in publicizing such a program through their own publicity channels. In any event, they should be notified and contacted to solicit their cooperation.

Attempts should be made to secure publicity in newspapers, including labor and neighborhood newspapers. Radio and television stations often have programs on which they make local announcements of this sort. Don't forget to send announcements to "Industrial Quality Control".

In general, mailing is handled by the local Education Committee and the secretarial services of members in the larger companies. It is desirable, however, to use professional mailing service as soon as the section is financially able to justify this move. Often local mailing organizations have mailing lists which are available and may be used for this purpose. Don't forget ASQC area lists as well. Chambers of Commerce and trade associations may have lists which can be purchased or rented for approved mailings of this sort.

6) Course Co-Sponsorship

It has been found extremely desirable to have ASQC Section Courses co-sponsored by local Chambers of Commerce, educational institutions, or industrial associations. Generally, the co-sponsoring group does little other than permit the use of its name as a co-sponsor. In specific instances, the co-sponsorship may go much beyond this to include the mailing of announcements to their members or to providing mailing list to the local sections. It may also

include financial assistance and other forms of area publicity. It can specifically help in obtaining publicity through local newspapers, radio and TV stations, bulletins, etc.

7) Questionnaires

Questionnaires can be extremely valuable both for the immediate and long-range planning of a training program. There are three types of questionnaire programs generally used. The first of these is a general information questionnaire circulated to the membership and perhaps to a general area mailing list. The purpose of this questionnaire is to get some idea of the interest of the potential registrants in the area. The following items might be subject matter for such a questionnaire:

1) Selection of date and time of course,
2) Selection of possible subject matter from a submitted list,
3) Occupational interests,
4) Educational background of those interested.

Another important questionnaire is usually given to enrollees at a training course at the first session. This includes such questions as:

1) Industrial position,
2) Educational background,
3) Field of activity,
4) Reasons for taking Quality Control course,
5) Source of information on the course being offered.

Upon completion of the course, another questionnaire can be of interest. This questionnaire would concern itself with:

1) Other courses or subjects that the enrollees would be interested in,
2) Comments on the arrangements for the present course,
3) The nature of the material presented (too technical, too simple, more illustrations, more explanations, etc.),
4) Areas of application that the individual might anticipate for the material presented,
5) Questions on Society participation.

8) Reports

It is extremely important that a detailed report on each course be prepared. This should include preparation, procedure, operational details, and financial details. The Saddoris Award Competition encourages this through the points given by the Committee on Education and Training for such reports. The file of these reports provide the basis for development and improvement of the educational programs of the individual sections, as well as for the National Committee on Education and Training. Much of the information contained in this paper is based on such section reports.

9) Certificates

It is desirable to give each registrant who attends practi-

cally all the sessions of an ASQC training course a "Certificate". Again, this is essentially psychological in its effect and is appreciated by the student and by his company which usually pays his registration fee. This certificate should merely state that the individual has attended a course covering a specific number of hours given by the ASQC Section at a certain time. This certificate should make no statement whatsoever as to the individual's qualifications in the field of quality control or related activities. In general, this certificate is signed by the President of the local section and the instructors.

10) Time-table

The following is a time-table that is typical of that necessary for the average course. This, of course, must be adjusted to the individual needs and activities of the section. It is offered here only as a rough guide to the major important planning schedule.

Item	Days Prior to Start of Course
Select general course content and instructors	90
Brochure, general outline	80
Prepare detail course outline	75
Finalize arrangements for room and instructional materials	72
Finalize brochure design	70
Finalize brochure printing arrangements and copy to printer	65
Start first mailing	45
Start second mailing -- if used	30
Complete mailing program	28
Follow-up on publicity to assure planned distribution	20
Check of arrangements for classroom and material	14
Check on arrangements for person to introduce course, handle late registrants, etc.	10

While the above outline may indicate areas requiring attention and some of the approaches that have been used successfully, it is necessary that each section course reflect the needs and interest of the local community. It is only through satisfying this local interest that section courses have proved to be such a successful activity of the Society. The three basic course purposes indicated at the beginning of this paper presuppose the need. For this reason, in closing, we should like to stress the fact that service to the local community must be an underlying premise of all the matters considered herein. It is our feeling that the success of the ASQC and of its activities is based upon the fact that it has well served the industrial community.

MULTIPLE COMPARISONS IN THE ANALYSIS OF VARIANCE

David L. Wallace
The University of Chicago

1. Introduction.

A common type of experimental data consists of measurements on the outcome of some process under several sets of conditions. For example, the yield of a chemical process is measured for several runs with each of six different catalysts. Or the knock rating of fuel is measured for samples from several lots from each of four suppliers. Several variable conditions may be studied at once as when impact strength is measured for ceramic insulators fired with each of five different types of glazes and assembled by each of five operators.

In experiments such as these, interest usually centers in the differential effects of the conditions: catalysts or suppliers or glaze types or operators. The effect of a condition is ordinarily interpreted as the average performance of the process under that condition over a long sequence of repetitions. A strong desideratum for the statistical analysis of such data is a method for comparing the effects of any two conditions, making proper allowance for the variability of the observed data. The purpose of this paper is to show how this need can easily be met, using a method due to Tukey (1). The problems and method are similar to those in the multiple comparisons of several treatments with a standard, presented to the 1955 convention by Dunnett (2).

2. The proposed multiple comparisons method.

Experimental data given by Davies (3) will be used to illustrate the method. During a filtering process in the chlorsulphonation of acetanilide, some of the product is lost. The percent loss was measured for three samples of each of five different blends of acetanilide and the data are shown below. (In the experiment, the rows corresponded to blocks of time in production, but this is ignored here to simplify the exposition.)

PERCENT LOSS OF PRODUCT

BLEND OF ACETANILIDE

	A	B	C	D	E
	16.9	18.2	17.0	15.1	18.3
	16.5	19.2	18.1	16.0	18.3
	17.5	17.1	17.3	17.8	19.8
Means	17.0	18.2	17.5	16.3	18.8
Variances	.25	1.10	.32	1.89	.75

The multiple comparison procedure requires the calculation of a quantity denoted by WSD (wholly significant difference), conveniently found in four steps.

1. The estimated variance s^2 of a single observation is found. Here it is the average of the five variances of the observations within blends and is .86 with ten degrees of freedom (DF). For k groups of n observations each, there will be $k(n-1)$ degrees of freedom.

2. The standard error of a difference of two means is found as $\sqrt{s^2(\frac{1}{n}+\frac{1}{n'})}$, here equal to .76 with $n=3$.

3. The LSD (least significant difference) is the product of the standard error of a difference of two means and the 5% (two-tailed) critical value of Student's t with the same degrees of freedom as s^2. A brief table of t values is given in table 1 and more values can be found in any statistics text. Here t is 2.23 and LSD = (.76)(2.23) = 1.7.

4. The factor $\left[\frac{WSD}{LSD}\right]$ for converting LSD to WSD is taken from table 1 according to the number of groups and degrees of freedom of s^2. Rough interpolation where necessary in this table is adequate. For five groups (blends) and ten degrees of freedom, the factor is 1.48 and WSD = (1.7)(1.48) = 2.5.

The WSD is then used to form multiple comparison statements. For any observed difference in blend means, the WSD is used as an "allowance" for variability. Thus a few of the possible conclusions are:

(i). The difference in percent loss between blend E and blend D lies in the interval 18.8 - 16.3 ± WSD or between 0.0 and 5.0.

(ii). The difference in percent loss between blend B and blend D lies in the interval 18.2 - 16.3 ± 2.5 or between -0.6 and 4.4.

The WSD was chosen so that all possible comparisons will be correct in 95% of such experiments. In about 5% of such experiments, one or more of the conclusions will be incorrect. This is called a 5% error rate experimentwise and will be discussed further in the next section.

More complicated comparisons than simple differences between blend means can be made if desired without increasing the experimentwise error rate. For example:

(iii). The difference between the mean of the percent losses for blends B and E and for D lies in the interval
$\frac{1}{2}(18.8 + 18.2) - 16.3 \pm 2.5$ or between -0.3 and 4.7.

(iv). The percent loss for blend D lies in the interval 16.3 ± 2.5 or between 13.8 and 18.8.

Conclusion (iii) follows from (i) and (ii) by averaging, but (iv) is entirely different--not strictly a comparison, Generally, if $c_1x_1 + c_2x_2 + \cdots + c_kx_k$ is some linear combination of the group means, then the proper "allowance" is the product of the WSD and the larger of (sum of the positive c_i) and (—sum of the negative c_i). In all examples above, this latter factor was one.

Recall the meaning of any conclusion, say (ii). The statement is a prediction with a 5% error rate experimentwise that in repeated use of this production process, the percent loss of product using blend B will average between 4.4 higher and -0.6 lower than the percent loss

using blend D. But the "allowance" allows for only those sources of variation occurring in the experiment. Thus if variation in percent loss between batches of acetanilide is greater than within batches, and if all of the samples in the experiment came from the same batch, then the predictions will not be adequate for new batches, and the true error rate if so used may be much higher than the nominal 5%.

The comparisons give a rather complete summary of the results of the experiment. One might be tempted to discard the entire set of conclusions and replace them by the simple one that there are no significant differences between blends. But this would be very wasteful. Blend D is surely as good (i.e. less percent loss) as blend B and may give as much as 5.0 less percent loss (which would be almost one-third less). Had the lower limit been 0.1 instead of 0.0, the difference would have been significantly different from zero, yet not appreciably more useful. The experiment may be too small to yield adequate information, but the multiple comparison intervals indicate not only whether a difference is significantly different from zero, but also how large it may be.

3. WSD, LSD and error rates.

If there had been only two blends, standard analysis would use Student's t distribution to set a confidence interval on the difference of the average effects. For two groups of n measurements each, the interval would be

$$\bar{x}_1 - \bar{x}_2 \pm t\sqrt{\frac{2s^2}{n}} = \bar{x}_1 - \bar{x}_2 \pm \text{LSD}$$

except that the LSD uses the estimate of variance based on all the groups. Each such statement is a 95% confidence interval.

All of the comparisons of section 2 could be made using LSD instead of WSD, and the intervals would all be narrower. But there is a serious objection to this procedure because of the large number of possible comparisons and the resultant large number of incorrect statements. That only a few of the possible comparisons would be made is not protection, since those few are chosen after looking at the data as most interesting. There are $\frac{1}{2}k(k-1)$ possible differences of k means--ten differences of five means. In a long sequence of experiments with five groups, the number of incorrect LSD statements would be about half the number of experiments and with more groups, the ratio would be higher. Because of dependence among comparisons in one experiment, several incorrect statements will often occur together in one experiment. Nevertheless, with five groups, using the LSD comparisons (and assuming the degrees of freedom to be large), 29% of experiments will contain at least one wrong statement; with ten groups, 67%; with three groups, 12 %. These last figures are called the error rate experimentwise--the percentage of experiments in which one or more statements are incorrect.

The WSD comparison system is based on the premise that the more stringent control of error rate experimentwise provides the user a more adequate set of conclusions than the looser LSD system with error rate per comparison. The $\boxed{\frac{\text{WSD}}{\text{LSD}}}$ of table 1 is just the necessary factor by

which the LSD must be expanded to reduce the error rate to 5% experimentwise. It is interesting to note how little expansion is needed to control the error rate and to see how it varies with the number of groups and with degrees of freedom.

4. Relation to the analysis of variance F-test.

The conventional analysis of variance procedure consists in decomposing the variance into several parts and carrying out one or more F-tests. The results of this procedure for the example are summarized below.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F-ratio
between blends	4	11.56	2.89	3.36
within blends	10	8.64	.86	

The F-ratio is slightly less than the 5% critical value 3.48 of F with 4 and 10 degrees of freedom, so that there are no significant differences among blends. This sole result of the F-test is approximately equivalent to one of the many conclusions available from the multiple comparisons statements and is certainly not an adequate analysis of the experimental data.

The analysis of variance decomposition serves a very useful function in multiple comparisons procedures in the calculation of the variance s^2 of a single observation. The denominator mean square of the F-test for between blend effects is exactly the estimated variance of a single observation calculated in step one in the calculation of WSD. This will be true generally and will be particularly useful in more complex sets of data.

5. Applications to more complex designs.

The WSD multiple comparisons procedure can be applied without difficulty to many sets of experimental data more complex than the single classification scheme with equal numbers of replications thus far considered. If $\bar{x}_1, \cdots, \bar{x}_k$ are the means for suppliers in an experiment involving suppliers, operators, machines, and days, and if each mean is based on the same number of observations, then provided that the variance s^2 and its degrees of freedom are obtained correctly, the procedures for getting allowances for differences of the supplier means is just as in section two. The correct variance s^2 is most easily found as the denominator mean square which would be used in the analysis of variance F-test for the main effects of suppliers. In this form, the necessary calculations are explained in most texts on applied statistics or on the design and analysis of experiments. Using the same experimental data, multiple comparison statements on the difference of machine effects (or operator effects) could also be made (using a different WSD, of course).

All analysis of variance procedures become more difficult when unequal numbers of replications occur for different conditions. With a single classification into groups of unequal size, a natural modification of the WSD procedure can be used. The variance s^2 can be obtained from the analysis of variance decomposition (or as the weighted average

of the within group variances). In step two, the standard error of the difference of means of groups of size n_1 and n_2 will be $\sqrt{s^2\frac{1}{n_1}+\frac{1}{n_2}}$ and will be different for each possible difference in means. For each difference of means, the LSD and WSD and allowance are found from the standard error exactly as with equal sized groups. The use of the $\left[\frac{WSD}{LSD}\right]$ factor does not give exactly a 5% error rate experimentwise for the resulting multiple comparison statements, but studies by Kurtz (4) indicate that the suggested procedure is a fairly good approximation. The actual error rate appears always to be somewhat less than 5%.

6. Shortcut methods.

When the data consists of a single classification into groups of equal size, short cut procedures using ranges of the measurements in a group as estimates of variability have been worked out. The sum of the within group ranges is multiplied by a tabled factor depending on the number and size of the groups and on the error rate to give directly an allowance for differences of group totals. An allowance for differences of group means, if that is preferred, is obtained by dividing by group size. For the numerical example of section two, the total range is $(1.0 + 2.1 + 1.1 + 2.7 + 1.5) = 8.4$ which gives, when multiplied by the critical factor .94, an allowance of 7.9 for differences of blend totals or an allowance of $(7.9)/3 = 2.6$ for differences of blend means. This compares closely with the allowance of 2.5 obtained by the more efficient method using variances. The allowances are used in exactly the same way.

Tables of the critical factors were given by Tukey (5) and can also be found in (6), (7), (8). (The entry for 5% error rate for four groups of size two is given incorrectly in (5) and (6) as 1.98. It should be 1.78.) Similar but slightly more complicated range methods for double classification into rows and columns were given by Tukey (5) (also in (6)) and modified by Kurtz et al (9) (also in (7)).

7. Other systems of multiple comparisons.

Several other systems of allowances for multiple comparisons statements have been proposed for use in problems like these. Each system has certain advantages and disadvantages. Dunnett (2) has constructed a system designed for problems in which one treatment or method is a standard or control. His system gives shorter allowances for the comparisons of each treatment with the standard at the cost of longer allowances for the less important comparisons of two new treatments.

The most important alternative system in which all differences of means are treated alike is that proposed by Scheffe (10). In constructing allowances for the differences of group means, it is the same as the factors corresponding to those of table 1 are all larger. This unalloyed disadvantage is sometimes compensated for by its superior mathematical manageability (e.g. for handling unequal sized groups) and by its shorter (than the WSD system) allowances for more complicated comparisons such as the comparison (iii) of section two or the regression coefficient of mean yield on temperature where the groups correspond to different temperatures.

For some purposes of experiments, none of these types of data analysis and summary are relevant. If the sole object of the experiment is to pick the best supplier, the analysis is easy: pick the supplier with the best mean. The real statistical problem is in designing the experiment to give a sufficiently accurate choice. Bechhofer (11) has treated this problem.

8. The table of factors.

Table 1 gives a factor for converting an LSD to a WSD with 5% error rate experimentwise for a selection of values of the number of groups and the degrees of freedom of the variance estimate. The first column gives 5% (two-tailed) critical values of Student's t. Other values of t can be found in any text or set of tables. Other values of the factor, adequate for practical use, can be obtained by rough interpolation in table 1, or alternatively from table 29 of Biometrika Tables (12) by the relation:

$$\left[\frac{WSD}{LSD}\right] = \frac{\text{Upper 5\% point of studentized range of k groups.}}{\sqrt{2}\ \text{(two-tailed 5\% point of t)}}$$

Table 1 was obtained as slight modifications of tables of Tukey (1) and Biometrika Tables (12).

Table 1.

Factor $\left[\frac{WSD}{LSD}\right]$ to convert an LSD to a WSD for multiple comparisons of k groups with 5% error rate experimentwise and with DF degrees of freedom for variance estimate. The first column contains two-tailed 5% critical values of Student's t.

		k, number of groups							
t	DF	3	4	5	6	8	10	15	20
2.45	6	1.27	1.42	1.54	1.63	1.77	1.88	2.06	2.19
2.31	8	1.25	1.40	1.50	1.58	1.72	1.82	1.99	2.11
2.23	10	1.24	1.38	1.48	1.56	1.68	1.78	1.94	2.06
2.13	15	1.23	1.36	1.45	1.53	1.64	1.73	1.88	1.98
2.09	20	1.22	1.35	1.44	1.51	1.62	1.70	1.84	1.94
2.02	40	1.21	1.33	1.42	1.48	1.59	1.66	1.79	1.88
1.96	∞	1.20	1.32	1.40	1.46	1.55	1.62	1.74	1.81

9. References.

(1) Tukey, J. W., "The problem of multiple comparisons," Duplicated manuscript, Princeton University, 1953.

(2) Dunnett, C. W. "Multiple comparisons with a standard," National Convention Transactions 1955, Ninth annual convention, American Society for Quality Control, p. 485-491.

(3) Davies, O. L. (Ed.) The Design and Analysis of Industrial Experiments, Oliver and Boyd, London, 1954.

(4) Kurtz, T. E., "An extension of a multiple comparisons procedure," Unpublished PhD thesis, Princeton University, 1956.

(5) Tukey, J. W. "Quick and dirty methods in statistics, part II. Simple analyses for standard designs," Quality Control Conference Papers, 1951, Fifth annual convention, American Society for Quality Control, p. 189-196.

(6) Duncan, A. J. Quality Control and Industrial Statistics, Richard D. Irwin, Inc., Chicago, 1952. Tables L1 and L2.

(7) Lindsay, G. (Ed.) Handbook of Social Psychology, Addison-Wesley Publishing Co., Cambridge, Mass., 1954. Chapter 8: Selected quantitative techniques.

(8) Federer, W. T., Experimental Design, The Macmillan Co., New York, 1955, Table II-4.

(9) Kurtz, T. E., Link, R. F., Tukey, J. W. and Wallace, D. L. "Some short cuts to allowances," Duplicated manuscript, Princeton University.

(10) Scheffe, H. "A method for judging all contrasts in the analysis of variance," Biometrika 40 (1953), 87-104.

(11) Bechhofer, R. E. "Multiple decision procedures for ranking means," National Convention Transactions 1955, Ninth annual convention, American Society for Quality Control, pp. 513-519.

(12) Biometrika Tables for Statisticians, Edited by E. S. Pearson and H. O. Hartley, Cambridge University Press, 1954.

STATISTICAL METHODS FOR ANALYZING PERFORMANCE VARIATIONS OF ELECTRONIC CIRCUITS

R. H. Hinrichs
Convair
A Division of General Dynamics Corporation

I. Introduction

Reliable operation of electronic equipment is recognized as a goal which is of great importance to our defense effort. However, the complexity of high-performance electronic equipment complicates the problem of achieving reliable operation. Attempts to reduce complexity may help up to a point, but we must obtain reliability in spite of complexity.

Reliability can not be obtained by testing and inspection alone. Reliability must be designed into the equipment. To accomplish this requires a well-rounded reliability program involving several functions. This paper considers only one of these functions: the assessment of parts compatibility.

It is first necessary to explain what we in the Convair Astronautics Reliability Group mean by the term "parts compatibility", because we use the term in an unusual way. We know that the performance characteristics of manufactured electronic devices vary because of variations in the electrical characteristics of parts used in their manufacture. If there is a high probability that the performance of the equipment remains within tolerance, when any combination of specified parts, randomly selected from stock bins, is used; then we say that parts compatibility has been achieved. On the other hand, if the equipment performance is out of tolerance, even though all of the parts are good and within tolerance, then we have an out-of-tolerance failure.

Note that this definition of parts compatibility relates only to out-of-tolerance failures. It does not relate to catastrophic failures, those which occur if an electronic part breaks, shorts, opens, or burns out. Non-occurrence of catastrophic failures, it should be emphasized, offers no assurance that out-of-tolerance failures will not occur.

Assurance of parts compatibility may be obtained by production testing, and, of course, such testing should always be done as "proof of the pudding." However, such measurement after the fact of manufacture is only proof of accomplishment. To prevent out-of-tolerance failures, something else is needed, and that is a method of assuring that parts compatibility has been achieved while the equipment is in the design phase, that is, prior to manufacture. Such assurance could be obtained by experimental methods. But for a complex piece of equipment, with dozens or hundreds of parts, an astronomical number of tests would be required to check a statistically significant number of combinations of parts.

What is needed is an analytical or mathematical method of assessing parts compatibility while the equipment is still in the design phase. In the remainder of this paper two such methods will be discussed. The methods make use of digital computers to handle the huge mass of calculations involved. The auxiliary operations involved will also be discussed.

II. Operations and Methods

In order to assure parts compatibility in a design, it is first necessary to know the nature of the variations in the parts characteristics themselves. That is, one must obtain statistical distributions for all of the parts. This is done at Convair by measuring all of the parts in Receiving Inspection. A copy of these data is sent to the Reliability Group. This procedure also assures elimination of the out-of-tolerance parts. The gathering and analysis of this parts data is an expensive project, but reliability can not be bought cheaply. It should be further pointed out that gathering the parts data is not a peculiar requirement of the particular methods chosen. Parts data are necessary to any method which is used to assure parts compatibility. This is because the statistical variations in the parts characteristics are the causes in a cause-and-effect relationship. The effect is the variation in the performance characteristic of the manufactured equipment.

At Convair, we make use of digital computers in the analysis of the parts data from Receiving Inspection. The data consist of measurements of the resistance of resistors, and the capacitance and sometimes the Q-factor of capacitors. The data are identified by standard part number and sent to the Digital Computing Laboratory. Here an IBM 650 computer is used to obtain the histogram and statistical parameters of each lot of parts. The statistical parameters used are the lot size, mean, standard deviation, and the first eight statistical central moments. (Ref. 1) These results and the tabulation of the histogram are printed out on a page and also punched out on IBM cards. The cards are kept in order to combine the results of several lots of the same part. Thus there is available a library of parts data, containing all of the information accumulated to date, on each standard part.

It is also necessary to have data available on the distribution of electron tube characteristics. This is obtained by running tube tests, and measuring all of the characteristics thought to be needed in further analyses. Here there are so many variables and levels of each variable that an orthogonal block design for the test is set up by a statistician, in order to do the testing most efficiently. The results of the test are subjected to the analysis of variance to determine the significant relations. Then curves are fitted to the tube characteristics by least-squares regression techniques. The IBM 650 computer is used for the analysis of variance, and the IBM 704 computer is used for the curve-fitting. Tube characteristic curves are obtained for static currents and for transconductance, amplification factor, and plate resistance. Also, equations and curves are obtained for the standard deviation, and third and fourth statistical central moments of the above variables.

All of this tube data and the parts data previously described is filed, and constitutes the basic information needed for the mathematical methods.

There are two such methods currently in use at Convair Astronautics, the Monte Carlo method and the method of moments. I will not go into the detailed mathematics involved in these methods, but I will describe their main features.

The Monte Carlo method utilizes the basic laws of probability to determine a result. Here we are trying to determine the statistical variation in electronic circuit performance from a knowledge of the statistical variation in the tube and parts characteristics.

The first step is to determine a mathematical relation between the circuit performance characteristic and the parts characteristics. This is expressed in the form of an equation or a set of equations. Circuit theory or network analysis is used to obtain the equations. From the equations, the circuit performance can be calculated for any combination of parts characteristics. In the Monte Carlo method, this calculation is made and repeated many times. Each time a different combination of parts characteristics is used. In order to obtain meaningful results, the numbers representing the parts characteristics must be chosen at random in such a way that they will have the same statistical distributions as the parts characteristics they represent. The parts and tube data, in the file previously described, are used to determine these statistical distributions.

In practice, both the generation of the random numbers and the calculation of the circuit performance from the set of equations are done by the IBM 704 computer. The result is a large quantity of numerical values of the circuit performance characteristic. A histogram of these values may be plotted, or an approximate statistical distribution may be obtained by any of several standard statistical methods. (Refs. 2 and 3) This is then the desired result.

In the method of moments, the statistical moments of the performance characteristic are obtained from the statistical moments of the parts characteristics. The moments are statistical parameters which have several important properties. For example, the first moment is the mean, or average value. The standard deviation is the square root of the second central moment. Higher-order moments characterize other features of a statistical distribution. Thus the moments have the important property that they are able to summarize, in a numerical way, the main features of a large body of data. It is also possible to construct an approximate statistical distribution, if several of the moments are known (Refs. 2 and 3).

The method of moments then consists of the following general steps:

1. An equation is written which expresses the circuit performance in terms of the parts characteristics, as was done in the Monte Carlo Method.

2. The equation is replaced by a Taylor's Series. For our purposes we can think of the Taylor's Series in this way:

   Each of the parts characteristics has an average value and a distribution of values about the average. Any individual part has a characteristic above or below the average value, and this is considered as a positive or negative deviation from the average value. Also, the performance characteristic deviates above and below its average value. The Taylor's Series is an equation which expresses the relation between the average value of the parts characteristics, the parts deviations, the average value of the performance characteristic, and the performance deviation. It would therefore be ideally suited for the present purpose, except for the fact that it is an approximation. However, for those problems for which the approximation is good enough, the Taylor's Series is convenient. It also has the tremendous advantage that it is a standard mathematical model; that is, it is a standard form of equation which may be used for a large variety of problems by merely putting in

different sets of numbers for different problems.

3. The next step in the method is to find the statistical central moments of the performance characteristic from the statistical central moments of the parts characteristics. Here the Taylor's Series is used, as well as a set of equations relating the central moments. These equations were derived especially for this purpose, and have been programmed and put on tape for use with a digital computer. This program tape is limited to fifteen parts or tube characteristics, but this is sufficient for most circuit applications. This is where one of the advantages of using a uniform mathematical model, such as a Taylor's Series, is realized: This same tape may be used for all problems, subject to the limitations previously mentioned.

4. After obtaining the central moments of the performance characteristic, its statistical distribution is obtained by any one of several standard statistical methods.

More detailed mathematical descriptions of the method of moments are given in two Convair reports, listed in the references (Refs. 4 and 5).

The advantage of the method of moments is its conciseness. Calculations are done only once, and a few numbers are obtained, which summarize a large amount of information. The disadvantages are the loss of detail in so summarizing, and the errors due to the various approximations which are made. So far, it is felt that the advantage outweighs the disadvantages.

III. Some Problems

The Monte Carlo method is presently being used to determine the variation in transient response of a fairly complicated circuit. The results were not ready in time for publication in these minutes. The Monte Carlo method is not new and no difficulty is expected. This problem is a good illustration of the main advantage of the Monte Carlo method: The method is convenient for problems which must be done entirely by numerical means, because they are too complicated to do by analytical methods.

The method of moments was first tried on a circuit which was designed solely to test the method. This circuit is shown in Figure 1. The performance characteristic analyzed was the phase shift of the transfer-function. Twenty of the circuits were built from twenty sets of randomly-selected parts and tubes. The parts and tube characteristics were measured and recorded, and their statistical distributions were calculated. The distribution of the resistance of the 560-ohm resistors is shown in Figure 2. It should be noted that this distribution is far from normal, although a sample of twenty is too small to judge the true shape. The distribution of the amplification factor of the twenty type 5814 WA tubes is shown in Figure 3. This looks more nearly normal, but is somewhat skewed.

The statistical moments of all of the parts were used to calculate the statistical moments of the phase shift. Then the distribution of the phase shift was calculated. The result is shown as the dashed curve of Figure 4. The phase shift of each circuit was also measured experimentally. The resulting distribution of the measurements is shown as the solid curve

of Figure 4. The experimental average value of the phase shift is about 5% greater than the calculated average value. This is believed to be due to neglecting the distributed capacitance of the wiring, when writing the original phase shift equation. To check this, a modified Monte Carlo method was used. The phase shift was calculated from the equation, by using the twenty sets of measured parts characteristics. The resulting distribution is shown as the dotted curve in Figure 4. This curve agrees closely with the one obtained by using the method of moments. This demonstrates an important fact: The method of moments is as good as the original equation, but great care must be taken in writing circuit equations to make sure that all effects are accounted for.

The calculated value of the standard deviation of the phase shift was about 8% greater than the experimental value. This was thought to be partly due to errors in the instruments used for measuring the parts and tube characteristics. Of course part of the errors are due to the various approximations involved in the methods.

Another problem in which the method of moments proved useful was that involving the frequencies of two radio beacons. This was a systems application of the method, rather than a circuit problem. The two uhf beacons in a missile were required to operate on the same frequency, or within a required tolerance. The antennas are connected to the beacons by coaxial cables. These cables are cut in the factory, but the tolerance on cable lengths is so wide that the beacon frequency may be anywhere within a considerable range. An approximate relation between the cable length and the beacon frequency was known. It was desired to find the statistical distribution of the difference in the two beacon frequencies. A close approximation to the desired result was obtained by using the method of moments. The cumulative probability distribution was calculated and is shown in Figure 5. It is too early is the program to obtain verification of the calculations. However, in this case, the analysis was useful in influencing a designer's decision. This problem is a case in which the simplicity of the method of moments was advantageous. In fact, the problem was so simple, that the numerical work was done on a desk calculator rather than by a digital computer. Had the problem been slightly more complicated, or had greater accuracy been required, a digital computer would have been used.

IV. Conclusions

Some general remarks should be made about the scope of the methods, and about their future use.

First, they are analytical or mathematical methods. Therefore their use is limited to problems which can be expressed by equations. However, because they are mathematical methods, their use is not restricted to electronic circuits and systems, but they should find use in the statistical analysis of many kinds of physical systems.

Second, the methods have so far been applied only to systems under ambient environmental conditions. This is considered as a preliminary step which is necessary to prove the methods. If the statistical variation of the performance characteristic is successfully calculated for ambient environmental conditions from parts data taken under ambient conditions; then it is believed that the same methods can be successfully used to calculate the statistical variation of the performance characteristics under service environmental conditions from parts data

taken under service environmental conditions.

Thus much work remains to be done, before we can have assurance of parts compatibility while equipments are still in the design phase. In carrying on this work, digital computers will be indispensable for performing the tremendous amount of calculations involved.

REFERENCES

1. H. Cramer, "Mathematical Methods of Statistics," 1946, Princeton University Press, Princeton.

2. J. T. Kenney and E. S. Keeping, "Mathematics of Statistics, Part Two," second edition, 1951, D. Van Nostrand Co. Inc., New York.

3. C. C. Craig, "A New Exposition and Chart for the Pearson System of Frequency Curves," Annals of Mathematical Statistics, Vol. VII, March 1936. Pp. 16-28.

4. R. H. Hinrichs "A Statistical Method for Analyzing the Performance Variation of Electronic Circuits," Convair Report No. ZX-7-009, Oct. 3, 1955.

5. R. H. Hinrichs, "A Second Statistical Method for Analyzing the Performance Variation of Electronic Circuits," by R. H. Hinrichs, Convair Report No. ZX-7-010, Feb. 15, 1956.

The above two reports may be obtained by addressing a request to:

Commander
Western Development Division Headquarters
Air Research and Development Command
Post Office Box 262
Inglewood, California

The request should be addressed to the attention of WDTC, and include the report title, number, date, contract number (AF-04(645)-4) and the Contractor (Convair Astronautics Division).

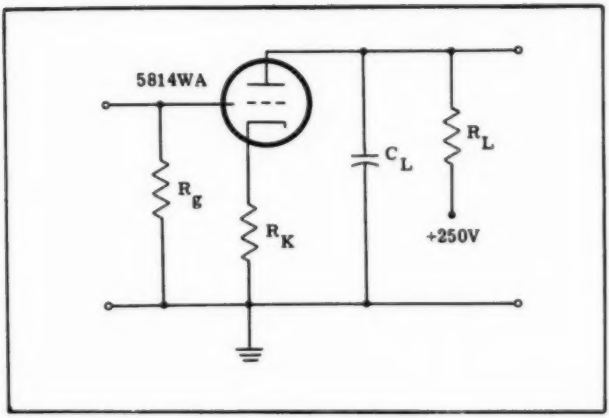


Figure 1. Simple Circuit

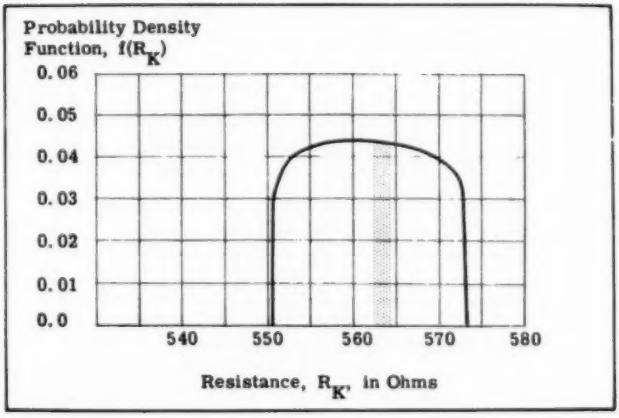


Figure 2. Probability Density Function

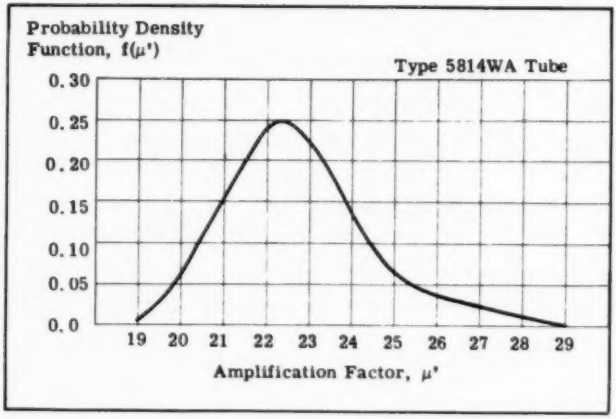


Figure 3. Probability Density Function of Amplification Factor

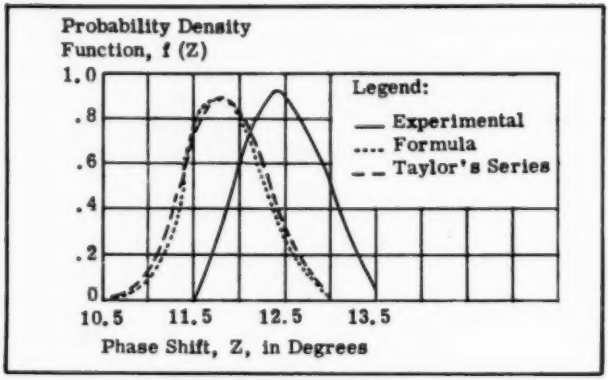


Figure 4. Comparison of Methods

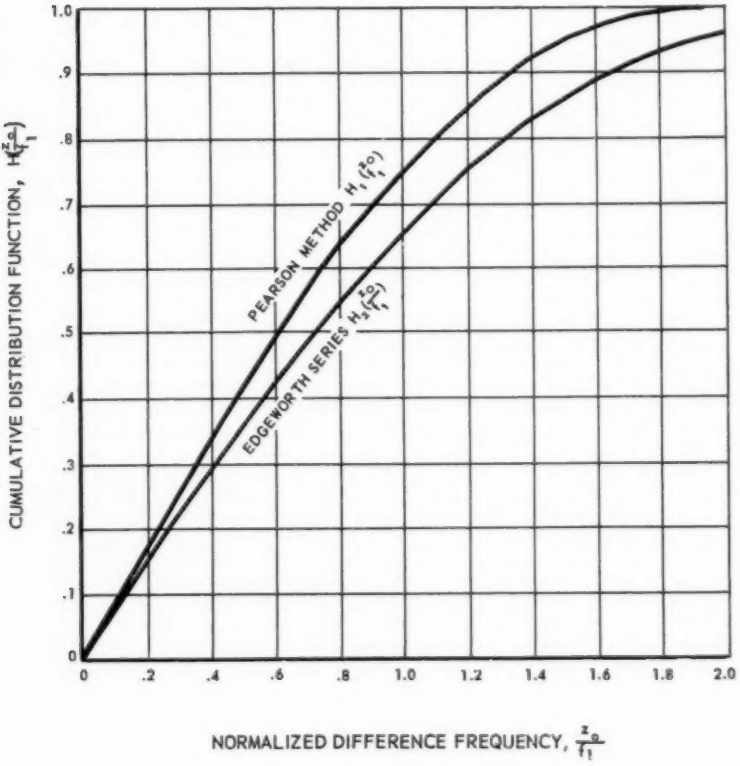


Figure 5. Probability Distribution of Beacon Difference Frequency

CONTROLLING THE QUALITY OF MANAGEMENT DECISIONS

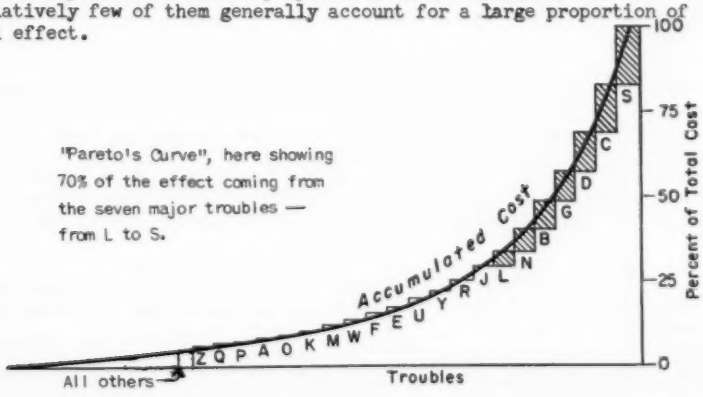
by Dorian Shainin
Rath & Strong, Inc.
Industrial Consultants
Boston, Massachusetts

In working with many industries we have come across three kinds of "quality" - each quite different. Quality of Conformance, which most SQC people strive to attain; Quality of Design, which is usually an engineering department function relating to grade of product (specifications) and to ingenuity of design; and Quality of Management Decisions - an area being touched by articles and books on Decision Theory and by Operations Researchers. Each kind of quality can be improved with the aid of certain statistical techniques. Perhaps some day several of the techniques in this last area may be collectively known as MQC as they assist in improving the Quality of Management Decisions.

Successful industrial management makes sound decisions based upon judgment and upon an ability to find facts. But facts, in the increasingly involved industrial situations today, are often elusive.

Modern industry invests rather heavily in the pursuit of facts. Research laboratories are continually busy in this chase. Engineers run tests on equipment and on part changes. Process people try new methods, personnel men follow up on aptitude tests to check the extent of correlations, and inspection foremen test lighting changes, alternate gages - the list in progressive plants can seem endless.

MQC uses three principles of imperfection in bringing facts to the surface. The first almost invariably shows the uneven (and thus most helpful) distribution of effect under management consideration versus the causes or sources of that effect. They often can be appropriately considered "the elements to be controlled". (1) It can be the simple plotting of each department's contribution to an important plant-wide function in a given period, such as total scrap or rework cost; each salesman's contribution to total sales; each vendor's trouble with returned parts; or the variation of each operator from standard time. If the effect is plotted as an accumulated value from the smallest contribution to the largest, a curve drawn through these points demonstrates this first principle of imperfection:- in any system of causes or sources of an effect, relatively few of them generally account for a large proportion of the total effect.


"Pareto's Curve", here showing 70% of the effect coming from the seven major troubles — from L to S.
Accumulated Cost
Z Q P A O K M W F E U Y R J L N B G D C S
All others
Troubles
100
75
50
25
0
Percent of Total Cost

Management thus has a "road map" to guide its attention for fruitful results in influencing the effect under consideration. For example if the effect is number of customer complaints (or schedule delays) working on relatively few of the causes or sources, say components of a complex assembly, can bring about a major reduction in the total effect.

Just as the first principle serves as a road map, there is a time when a street directory helps as management nears its destination. The second principle of imperfection draws upon a fact which made SQC what it is today: repetitive activity inherently displays variable results. A reduction of the spread of such variation and/or a change in its average level is often an object of management control. It can generally be brought about by examining a breakdown of the total variation into certain components. The second principle of imperfection states that when this is done, one part or component usually is found to be much larger than the others. (2) It provides a useful clue to aid in the detection of the cause of that component. The successful corrective action will be correlated with some element of that type of variation.

A company wanted to know what factors had a major effect on the sales of their product, used in the construction of homes. Its sales executives honestly felt there were innumerable variables involved, but the major ones were undoubtedly such things as housing starts, disposable income, and the amount of money spent on national advertising.

A trade association distributed monthly the combined sales figures of all member companies, state by state and nationally. It was a straight forward job to plot these components of total variation, the year by year, month by month, and state by state differences. The average sales one year to the next were almost the same. Month to month showed a mild seasonal effect, but not large. State to state differences were ten to twenty times as great.

The sales executives' reactions to this plotting were interesting. While they now conceded the three factors they originally suspected should have shown up as year to year variation, they felt the revelation of this new approach could be readily explained; state to state differences were obviously due to such things as population, climate, and industrial activity differences.

These could each be checked out by correlation plots. If the plot of sales, on the vertical scale, was influenced to a major extent by the suspected factor, on the horizontal axis, then the scatter of points about a trend line would be small - on the assumption that all other factors are varying at random while that on the horizontal scale remains fixed at a value. This situation can be controlled in certain designed experiments, and needs to be checked or objectively considered in a case like the one being discussed to assess the validity of such a conclusion.

Rather unexpectedly, plots of population, climate, and industrial activity versus sales each showed a wide scatter. Some other factor, varying at random while each of these is fixed, must be having a major influence on sales!

This conclusion encouraged three weeks of plotting of other factors

correlated with state to state differences. Finally one factor did show narrow scatter against sales, but in two, non-parallel bands. Another factor was interacting with the important one! Another week of plotting uncovered that one also. Now that company knows how to proceed to influence sales of its product favorably, and most effectively.

The third principle of imperfection applies to situations having several causes of variation acting at the same time and in such a way as to make a physical separation of effects from each cause virtually impossible. The principle states that it is unlikely in such situations that the important causes are simultaneously at their optimum levels with regard to desirable results, if they are controlled at all. Two devices are employed in the process of conducting tests to determine the facts, which enable management to see what causes are important and what are the optimum levels for them. One is called balancing, and the other randomizing. (3)

Figure 1 is of a layout or plan for an experiment to study the effect of two factors, A (advertising coverage) and B (type of advertising) on sales. A_1, A_2, and A_3 are three levels of coverage, say, low, medium, and extensive. B_1 and B_2 stand for two different types of advertising.

	A_1	A_2	A_3
B_1	trial No. 1 2	5 6	9 10
B_2	3 4	7 8	11 12

Figure 1

Each trial would represent a month's sales during which the particular combination of coverage and type, intersecting in that particular square of Figure 1, would be run. Since the sales for this product are affected, if at all, with a short or no lag time after the appearance of the advertisement, each trial would represent a month, spaced a month or two weeks apart.

If the tests were run in the same sequence as the trial numbers, it would be possible for other variables that change with time to "confound" with A and B and give an incorrect cause and effect relationship. Such variables might be changes in products of competitors, shifts in their strategy, and changes in consumer tastes or income. But if the 12 trial months are run in <u>random</u> sequence, say, 11, 6, 3, 10, 1, 8, 9, 4, 12, 5, 2, and 7, then <u>all</u> other variables changing with time will cause a lack of repeatability among pairs of results in the same square. These differences within pairs will determine the size of a single statistically determined number for experimental or background error.

Suppose the sales figures recorded when the test was run were:

	A_1	A_2	A_3	AVERAGES
B_1	190 210	235 265	290 270	243.3
B_2	175 165	235 205	240 260	213.3
AVERAGES	185	235	265	

The differences among the average when A_1, A_2 and then A_3 were present are 50 from A_1 to A_2, and 30 from A_2 to A_3. These differences cannot be affected by the effect of B_1 versus B_2, since both B's were equally represented in each of the averages for the A's. The effect of B's has been balanced out, for this part of the analysis. So the difference of 50 and 30 can only show the effect of any real differences caused by changing A_1 to A_2 to A_3 plus the chance effect of experimental error.

Six figures contribute to the experimental error, the differences between pairs of results in the same square 20, 10, 30, 30, 20 and 20 giving an average error of 21 2/3. When this figure is divided by the square root of the number of readings going into each average for A we get roughly the expected contribution of the error to differences among the averages for the A's. That gives us 10 5/6, not as much as the observed differences of 50 and 30. (4)

The differences between the averages when B_1 was present and when B_2 was present is 30. The expected contribution to this difference from the error is 21 2/3 divided by the square root of 6, or about 8.8. The effect of A's has been balanced out for this part of the analysis.

The statistical computations for a formal analysis of variance (not those done here) result in a numerical value of probability that the error can cause the observed differences among the A averages and the B averages. When the probability is low, the effect of A or B is called significant or real; when it is high, not significant. This table of odds shows the interpretations usually given to these results.

Odds	Decision
Less than 20 to 1	Not Significant
Less than 100 to 1	Questionably Significant
Greater than 100 to 1	Significant
Greater than 1000 to 1	Very Significant

Read as "odds of less than 20 to 1" that the result could be caused by other factors (error) oftener than one time in 20; "odds of greater than 100 to 1" that the difference could be caused by other factors (error) less often than one time in 100.

Experimental designs permit decisions, with a high and known probability that they are correct, as to whether real cause and effect relationships exist. Many factors can be incorporated in the same experimental design. So it is possible to determine which factors should be

controlled and at what levels they should be kept to get the desired results.

When members of management are exposed to such as these Management Quality Control (MQC) techniques in their own work, their use gradually becomes "second nature". Then they always aid good judgment so that better decisions are made in important areas.

Foot Notes:

(1) "Universals in Management Planning and Controlling", J. M. Juran, THE MANAGEMENT REVIEW, NOVEMBER, 1954.

(2) Due to L. A. Seder, the "Multi-Vari Chart" concept.

(3) Due to R. A. Fisher, the concepts of design of experiments and analysis of variance.

(4) This approach needed the word "roughly" in the previous sentence because it is purposely oversimplified to bring out more clearly the type of relationship involved. And we can further say, while the expected error is roughly 10 5/6, the actual contribution of error in this experiment could have been larger. And contributions further away from the expected have a smaller probability of occurring.

QUALITY CONTROL - ITS RELATIONSHIP TO RELIABILITY

Clair A. Peterson, Colonel, USAF
Headquarters, Air Materiel Command

Most of you are of course aware of the great emphasis that has been placed within recent months on the subject of reliability - particularly with reference to electronic equipment, which today comprises a major and vital part of our modern weapon systems.

Both industry and the military services have organized committees and task groups, concerned with the various factors affecting reliability; in order to determine, within their particular fields of interest, ways and means of improving reliability. Among the numerous magazine articles published, and the many addresses delivered in connection with such committee activity, it is interesting to note the frequent references made to quality control, and the important part they acknowledge it must play in any reliability program. Unfortunately, however, many of these articles and addresses reveal varied and mixed opinions as to the precise part quality control, as a function, should play in the overall reliability program.

To be sure, this function is most important; but it must be remembered that there are certain boundaries within which quality control must work; since it is essentially a supporting role, rather than one of direct responsibility. To further qualify this statement, it is first necessary that we understand the difference between "quality control" as a function, and "quality" as a characteristic.

Generally speaking, we consider quality control to be a conformance function; that is, conformance to certain quality standards which have already been established. "Quality Control," then, or control of quality, is construed to mean a system of controls which will assure that the quality of the end product will be as specified. Quality, on the other hand, must be expressed in terms of certain physical characteristics designed to obtain mission objectives.

Reliability, being dependent on quality, can be achieved only when adequate controls are exercised over quality. This means, then, that the quality control program, in coordination with the engineering, production planning, and other elements, must be designed precisely around the quality and the reliability requirements of the particular equipment under consideration.

Once the designer has finished his job, he must then look to quality control to take over the job of seeing that design requirements are being met; and also seeing that he is kept informed on any conditions which would indicate inadequacy of design or of test requirements, or conditions which indicate places where certain improvements could be made.

The effect, then, that reliability will have on quality control is this: that quality control can no longer be an independent and self-sufficient function, but instead must initiate its planning in

consonance with all other planning; that is it must develop a day-to-day working relationship with engineering, purchasing, production, and management; it must produce factual data and statistics - not only for its own use, but for the use of other elements of the organization concerned with reliability; it must hold a position co-equal to other executive positions within the organizational structure; it must staff its organization with talents and skills commensurate with the technical complexities of the product; and it must maintain a progressive and continuing program of training, concurrent with technological advancements - both in respect to the product and to quality control techniques.

As the trend toward placing on the contractor more responsibility for design and reliability progresses, likewise the quality control responsibility increases. Quality control must team up with the engineer in the development of test programs, test apparatus, and test facilities. Specifications should be reviewed for adequacy, subcontracting programs and purchase documents should be reviewed, and necessary liaison should be established between prime and subcontractor. The quality control effort must now be extended into every area where quality and reliability are affected.

Perhaps we should here consider what these areas may encompass, and also consider what Quality Control within the Air Force visualizes as a total program for reliability.

So far our discussion has referred to quality control only in the design and production stage; however, once the desired reliability has been obtained, it must be maintained - or the whole effort is lost.

Then we must recognize the need for a program which continues through the whole life cycle of the equipment - commonly referred to as the "cradle-to-the-grave" concept.

We at Headquarters AMC have already developed quality control programs in our supply and maintenance depots, and we are therefore in a position to introduce certain reliability assurance procedures in these areas.

These procedures will no doubt stem from and will be correlated to the reliability requirements established for the new equipment; utilizing the quality standards, test procedures, and acceptance criteria developed during the period of development and production of the original equipment. Through this process it should be possible for quality control to assure that the reliability built in during the procurement phase will be continued through the storage and maintenance phase.

One phase which at present has not been adequately covered in the life cycle quality control concept is the quality of maintenance while the equipment is in the hands of our operating commands. In order to assure that the original reliability built in during manufacture is continued during depot storage or maintenance, it is necessary that a more formalized quality control program be established at the operating

bases. Here again it will be necessary to incorporate specific procedures for checking reliability characteristics derived from requirements established for the particular equipment under consideration.

It can be expected that future contracts for weapon systems may require contractors to furnish much more comprehensive technical information for the operation, maintenance, and testing procedures, related to the reliability requirements for the equipment they manufacture. This information would then be useful to quality control in developing the life cycle quality control program for each weapon system.

AN OPERATIONAL RESEARCH STUDY ON SHIPPING TANKAGE

J. C. Dickson
Humble Oil & Refining Company

Introduction

Shipping large amounts of oil or oil products by tankers gives a refiner serious problems. The tankers do not arrive at fixed intervals and remove a fixed amount of oil. Instead, the time between tankers varies and forces the refiner to build shipping tanks. These tanks are insurance against delaying tankers or varying the production of oil.

And complications set in when the demand for the product is seasonal and the tanks used for seasonal storage are used to help out the shipping tanks. This is, of course, an advantage to the refinery but finding the best amount of shipping tankage gets rough.

This was the problem tackled by our Operations Research team; the team at first was made up of a physicist, a man familiar with existing tanks and lines, a man familiar with overall refinery economics, and two men from an Operations Research consulting firm.

It is the purpose of this paper to describe how we attacked this problem and the conclusions we reached.

Description of the Problem

As usual with Operations Research work, the first duty was deciding what needed to be done. There was little doubt the problem was one of simply minimizing costs and that employee morale, sales policies, etc., did not enter into the picture. So we could separate the costs which were pertinent to this problem:

(1) Cost of delaying tankers
(2) The operating cost on shipping tanks plus capital charges on these tanks
(3) The cost of production cutbacks
(4) The cost of pumping oil to or from seasonal storage
(5) The capital charges on oil in inventory

Next we needed to describe the physical system. Figure 1 is a rough outline of the model we used; the symbols used are defined in Table I.

The demand for oil on any day, D_i, is a function of the oil in tanks, planned production, and a random variable. This last thing, random variable, is used to represent the randomness in the arrival of tankers. The production on day i, P_i, is a function of the oil in the shipping tanks and the oil to be transferred to seasonal tanks. If the shipping tanks are too full, the production must be decreased.

The oil transferred to seasonal tanks on day i, T_i, depends on the amount of oil in the shipping tanks, the amount of oil in the seasonal tanks, and the day of the season. The day of the season is here because the seasonal tanks are used for other products part of the season. Once

these functions are defined, material balances can give us the inventory on each successive day and the delay time on the tankers. This is all we need to calculate the operating costs. Figures 2, 3, and 4 present the complete functions for P_i, T_i, and D_i.

Daily Production

Figure 2 shows five modes of production which we take into account. I will not try to explain all of these, for two examples should give the pattern.

The first mode is $P_i = P_p$. In other words the production is equal to the average if two conditions are met:

(1) $C_R < I_{Fi} < C_c$; which says that the inventory in the shipping tanks must be within certain limits, and

(2) $I_{Fi} + P_p - T_i \leq F$; which says that the planned production will not overflow the tanks after transferring oil to the seasonal tanks.

The fifth mode is $P_i = iP_p - \sum_{j=0}^{j=i-1} P_j$, which says the production is the planned production up to and including day i less the actual production up to and including the day before. This is the right mode if $iP_p - \sum_{j=0}^{j=i-1} P_j \leq P_p + \Delta_2$; i.e., if the production calculated by this mode is no more than some amount, Δ_2, over the average production.

Daily Transfer

Figure 3 shows the five modes of operation for the transfer. As in the case of P_i, only two of the modes will be explained.

The first mode says $T_i = T_{max}$. This is the right mode if

(1) $I_{Fi} \geq C_u$; i.e., the inventory is above a certain limit, and

(2) $S'_i \geq T_{max} + T_{si}$; i.e., pumping T_{max} barrels into tanks having T_{si} barrels in them will not make the total inventory greater than the allowable S'_i.

The fifth mode says $T_i = 0$. So no oil is transferred if

(1) $C_L < I_{Fi} < C_u$; i.e., the shipping inventory is within certain limits, and

(2) $I_{si} \leq S'_i$; the inventory in seasonal storage is not greater than allowable.

Daily Demand

Figure 4 shows the form of the daily demand for oil. Here is how we decided upon this.

(1) The product was picked up by two types of tankers. The first type takes on its main load of some other product and uses the product

under study for a filler. We called these "small cargo tankers." The second type is the one whose main cargo is the product under study. We called these "large cargo tankers."

(2) Other studies on shipping systems have successfully used Poisson distributions to describe the frequency of tanker arrivals. Our records on tanker arrival looked like they fit a Poisson distribution. (We later found that this had its limitations!)

(3) Some scheduling of tankers does take place. The guiding light in this scheduling seemed to be the difference between actual inventories and planned inventories. To account for the lag time in the exchange of inventory information between the "schedulers" and the refinery, a scheduling interval was selected and the average arrival rate of tankers was changed only at this fixed interval.

(4) Only large cargo tankers are scheduled.

These considerations led to the equations shown in Figure 4. The average number of small cargo tankers arriving on any day and their average cargo size was known. To calculate the demand represented by these tankers, we simply selected a number at random from a Poisson distribution to represent the number of these tankers. The average of this distribution was equal to the average number of small cargo tankers arriving per day. We then sampled the distribution of cargo sizes and got a size for each of these tankers. The sum of these cargoes is the demand for oil by small cargoes.

The tankers which had large cargoes are scheduled. So we had to have a method for deciding on the average or expected number of large cargo tankers arriving. The equation in the middle of Figure 4 is this adjustment. The desired shipment of oil for a short period is equal to the planned production, plus any inventory above planned, plus any cutbacks in production that have accumulated. When divided by the average cargo size for the large tankers we get the expected tanker arrival rate. The number of tankers and their cargo sizes is selected just as they were for the small cargoes. The total demand is the demand of the small cargoes plus the demand of the large.

The term "J" which is in the formula for M_L is the planned inventory in seasonal storage. This is calculated from the idea that the inventory in this storage should be reduced at a constant rate from the first day of the season to the last day that this tankage is available.

I will not go into the details of selecting random samples.

The Analysis of the Model

Now that all of the functions had been defined, the only job left was finding the best values for the volume of shipping tankage and the control limits, C_c, C_u, and C_L. The nature of the model cried for a Monte Carlo solution. This type of solution is hard to describe briefly but I will attempt to do so anyway.

In systems which include randomness or unpredictability, true analytical solutions are usually impossible. The Monte Carlo technique for solving these equations is similar to the way you would decide the odds for winning at solitaire - you would play the game and see how often

you win. Of course, with a large bulky "game" like a shipping system you cannot actually play very many times; so, a model is constructed which is analogous to the shipping system and which can be "played" rapidly. We make the assumption that what is good for the model is good for the system.

We tried various values for the things we were interested in and used our model to estimate the operating costs. By playing the game for a number of seasons and averaging the operating costs, we estimate the average operating cost of the shipping system. The more seasons we play the game, the better our average.

At first we planned to calculate the possible error in this average after each season and repeat the calculations until this error was reduced to insignificance. It did not take long to discover that too many seasons would be needed. Consequently, we had to find out how to live with a large error or how to reduce it some other way. Actually, we did a little of both.

It turned out that the large error was caused by the way we selected the number of tanker arrivals. The possibility of an extreme number of tankers arriving on any given day had been blocked by simply limiting the maximum possible value in our Poisson distribution for the tanker arrivals. On the other hand, there was no limit to the number of consecutive days on which no tankers arrived. The chance of a long run without tankers coming in was very small for a given season but when you are working with a sample equivalent to 40 or 50 years, the chance of this happening somewhere in the sample is significant. It happened, then, that occasionally one of our seasons would contain a successive run of nine or ten days without any large tankers arriving. The cost of the production cutbacks which had to follow this overshadowed everything else. Please notice that this type of risk can occur in most Monte Carlo problems.

We improved this a little bit by reducing the scheduling interval but we still had a large error in our final cost figures. The question then was whether or not we could live with this error.

A review of the entire study was made and it was found that the tankage for this shipping system would be changed by large increments to reach any optimum. In other words, 50,000 or 100,000 barrel tanks would be bought. A 10,000 barrel change has little meaning. Therefore, we didn't need the <u>exact</u> value for the optimum tankage. This made the error easier to live with. In fact, when considering this, the expected error of a 40-season sample was small enough to let us select the best tankage.

This error did not cause too much grief with regard to the decision limits either. An analysis of the possible effects of changing these limits showed that there would be a broad region of essentially no effect bounded by a region of extreme effects. For example, if C_L, the decision limit used for transferring oil from seasonal storage, is gradually raised above zero tanker waiting costs drop. As this limit is raised more and more it stops affecting the tanker waiting costs. Further, it does not show any other effect until it approaches the limit of the shipping tankage where the cost of production cutbacks jumps. All of this gives a broad area of equally good values for this decision

limit. The other decision limits can be shown to behave in a like manner.

So we found that we could live with the large error in the average operating cost, and recommendations regarding shipping tankage and the decision limits have been presented to management.

General Remarks About the Problem

This study had another purpose - providing us with information about shipping systems in general and the pitfalls in other tankage studies. It is my opinion that this project was successful in this respect. We will never build another random tanker arrival model without limits to both the number of tankers arriving in a given day and the number of days without a tanker arrival. If this had been included in our model for this study, the manpower spent in analyzing it might have been cut in half.

In future studies on shipping systems the model would not be without an "anticipator mechanism." We know that in the actual shipping system the exact dates of tanker arrivals are known two or three days in advance. Being able to predict trouble you can often prevent it. This can make a big difference in the amount of insurance you need. It may be hard to find the nature of the prediction and the methods of prevention but it is worth looking for.

At one time we hoped that this model would serve as a general model which we could adapt to any shipping system. This is not so. Future studies of shipping seasons will require new models; and if we don't try to make it into a general model, the effort and cost of solving the problem will be less.

TABLE I

DEFINITION OF SYMBOLS

i = day of season being worked
D_i = demand for product, day i
P_i = production of product, day i
T_i = transfer of product from shipping tanks to seasonal tanks, day i
I_{Fi} = inventory in shipping tanks, day i
I_{Si} = inventory in seasonal tanks, day i
S'_i = maximum allowable inventory in seasonal tanks, day i
C_{Lij} = cargo size of the jth large tanker, day i
C_{sij} = cargo size of the jth small tanker, day i
N_{Li} = number of large tankers, day i
N_{si} = number of small tankers, day i
M_L = expected number of large tankers
M_s = expected number of small tankers
μ_L = average cargo of large tankers
μ_s = average cargo of small tankers
T_{max} = maximum allowable transfer rate
P_{max} = maximum production
P_p = planned production or desired average production
C_c = upper decision limit for production cutback
C_u = upper decision limit for transfer to seasonal storage
C_L = lower decision limit for transfer from storage
C_R = lower decision limit for production increase
$\Delta 1$ = volume of low profit product being produced at P_p
$\Delta 2$ = volume of highest profit production not produced at P_p
F = capacity of shipping tanks
S = capacity of seasonal tanks
N = seasonal storage can be used only the first N days of season
r = scheduling interval
E_i = a random variable

FIGURE 1

THE SHIPPING SYSTEM

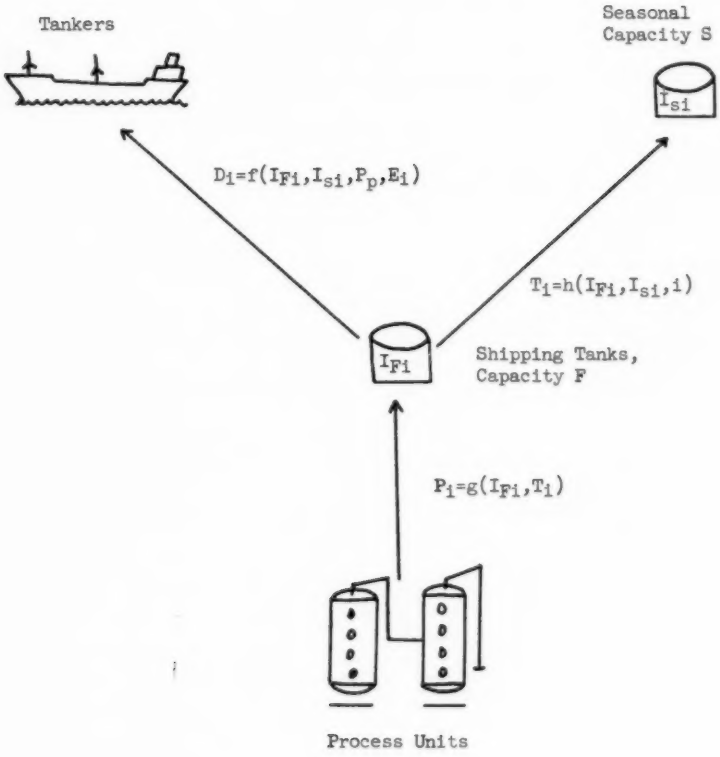


FIGURE 2

DAILY PRODUCTION, P_i

$P_i = P_p$ when $C_R < I_F < C_c$ and $I_{Fi} + P_p - T_i \leq F$

$P_i = P_p - \Delta_1$ when $I_{Fi} \geq C_c$ and $I_{Fi} + P_p - \Delta_1 - T_i \leq F$

$P_i = F - I_{Fi}$ when $I_{Fi} + P_p - T_i > F$

$P_i = P_p + \Delta_2$ when $I_{Fi} < C_R$ and $iP_p > \sum_{j=0}^{j=i-1} P_j + \Delta_2 + P_p$

$P_i = iP_p - \sum_{j=0}^{j=i-1} P_j$ when $iP_p - \sum_{j=0}^{j=i-1} P_j \leq P_p + \Delta_2$

FIGURE 3

DAILY TRANSFER, T_i

$T_i = T_{max}$ when $I_{Fi} \geq C_u$ and $S'_i \geq T_{max} + I_{Si}$

$T_i = S' - I_{Si}$ when $I_{Fi} \geq C_u$ and $T_{max} > S' - I_{Si}$
or $I_{Fi} < C_u$ and $I_{Si} \geq S'$

$T_i = I_{Si}$ when $I_{Fi} \leq C_L$ and $I_{Si} < T_{max}$

$T_i = -T_{max}$ when $I_{Fi} \leq C_L$ and $I_{Si} > T_{max}$

$T_i = 0$ when $C_L < I_{Fi} < C_u$ and $I_{Si} \leq S'$

$S' = S$ when $i \leq M$

$S' = S - T_{max}(i-M)$ when $M\ i \leq N$

$S' = 0$ when $i > N$

$$M = N - \frac{S}{T_{max}}$$

FIGURE 4

DAILY DEMAND, D_i

$$D_i = \sum^{j} C_{Lij} + \sum^{j} C_{sij}$$

where $\sum^{j} C_{Lij}$ is the sum of a sample, of size N_{Li}, selected at random from the distribution shown in Figure 5. N_{Li} is selected at random from a Poisson distribution having a mean of M_L.

$$M_L = \frac{I_{Fi} - 1/2\ F + I_{Si} - J + (i-1)P_p - \sum_{j=0}^{j=i-1} P_j}{r} + P_p - \mu_s M_s$$

$$\mu_L$$

$J = (i+r)\frac{S}{N}$ when $i + r \leq N$

$J = 0$ when $i + r \geq N$

M_L is calculated only if i is an exact multiple of r.

$\sum^{j} C_{sij}$ is the sum of a sample, of size N_{Si}, selected at random from the distribution shown in Figure 6. N_{Si} is selected at random from a Poisson distribution having a mean of M_s.

Figure 5

CUMULATIVE DISTRIBUTION FOR LARGE CARGOES

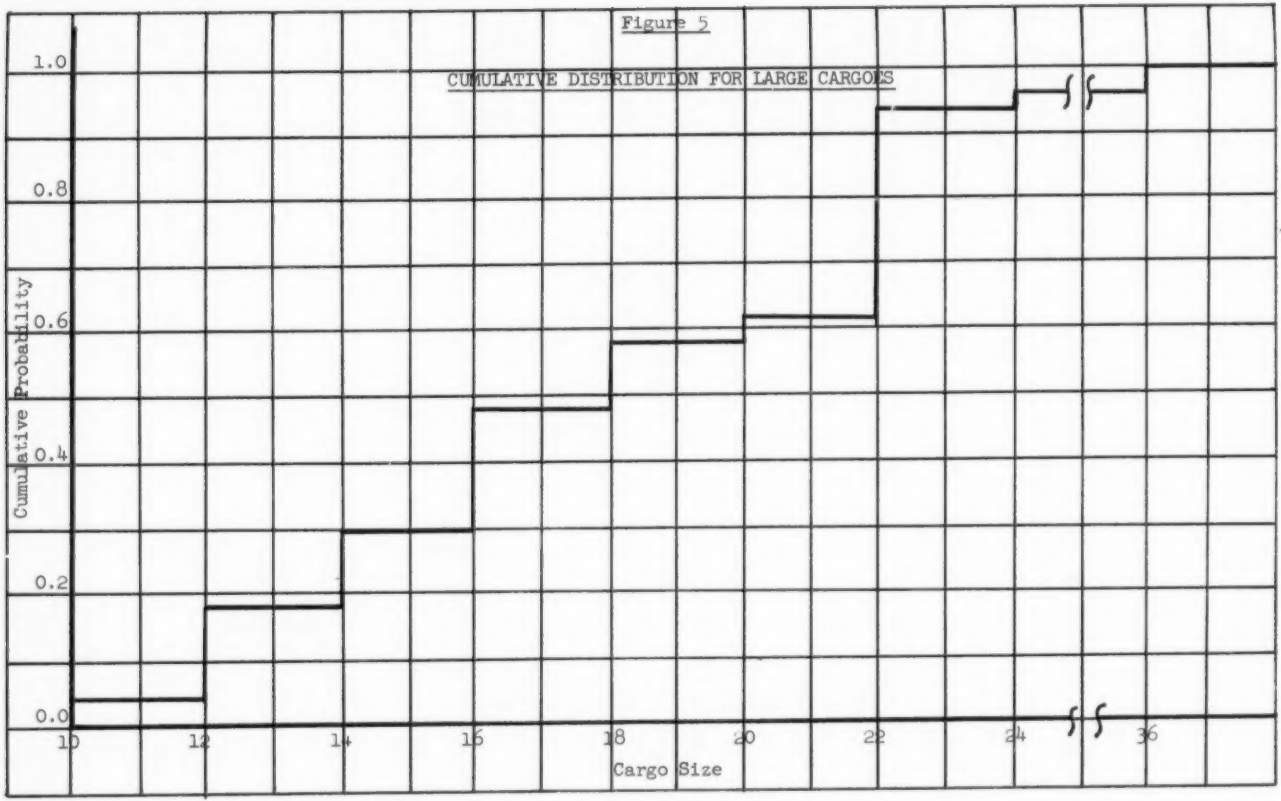
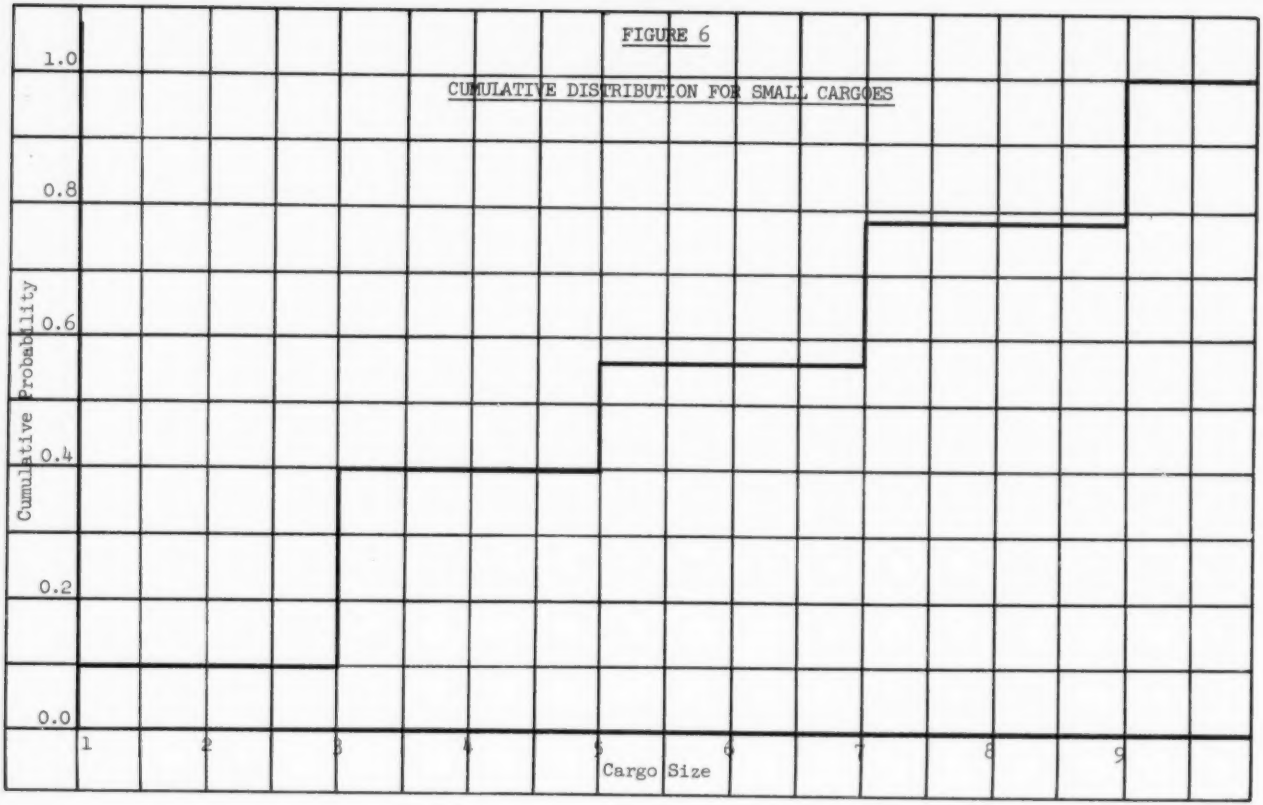


FIGURE 6

CUMULATIVE DISTRIBUTION FOR SMALL CARGOES



THE DIGITAL COMPUTER AND ITS APPLICATIONS

Donald E. Hart
General Motors Research Staff

You have all heard digital computers referred to as "giant brains" or "thinking machines." I should like to dispel the aura of science fiction which surrounds these machines by showing in some detail how we go about using a computer to solve a problem. I shall draw an analogy to a computer with which you all have some familiarity—an ordinary desk calculator.

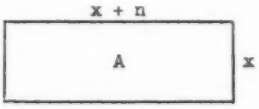
If we had a desk calculator and wished to solve the same problem over and over again, we would undoubtedly equip this desk calculator with an operator. Now, let us assume that this operator is almost a complete idiot, but that we have been able to teach this operator to recognize a limited set of written instructions. These are very simple instructions, but if we write down enough of them in the right sequence, we are able to use this computer to solve large and complex problems. If we present a sequence of these instructions to our operator, he will carry them out rapidly and without error; however, if we give him an instruction which is not in his repertoire, he will become confused and stop working. In order to show how we would get this computer to solve a very simple problem, I shall go through the seven steps that we must always go through in solving any problem on a digital computer.

I. Mathematical Formulation
II. Mathematical Analysis
III. Programming
IV. Coding
V. Checkout
VI. Production Computation
VII. Evaluation of Results

THE PROBLEM

Given: A rectangle with area = A sq. in.
One side is (n) inches longer than the other.

Find: The lengths of the sides.



STEP I - MATHEMATICAL FORMULATION

The problem must be described in mathematical language.

$$x(x + n) = A$$

$$x^2 + nx - A = 0$$

This is recognized to be of the form of a quadratic. Since a certain amount of effort is to be expended in preparing this problem for computer solution, it is desirable to set up a routine which will solve any quadratic equation:

$$ax^2 + bx + c = 0$$

Where, for our special case:

$$a = 1$$
$$b = n$$
$$c = -A$$

STEP II - MATHEMATICAL ANALYSIS

This step is necessary in order to reduce the mathematical formulas to a numerical form suitable for machine solution.

For the quadratic,

$$ax^2 + bx + c = 0$$

if numerical values for a, b, and c are given, the values of x cannot be found directly.

If the quadratic can be manipulated into the form,

$$x^2 + 2rx + r^2 = (x + r)^2$$

the square root can be extracted leaving only the first power of x. When this is done, the following familiar expressions are found.

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

$$x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

By plugging in numerical values of a, b, and c, it is possible to find x_1 and x_2. Unfortunately, a computer doesn't have "plug into" as one of its instructions.

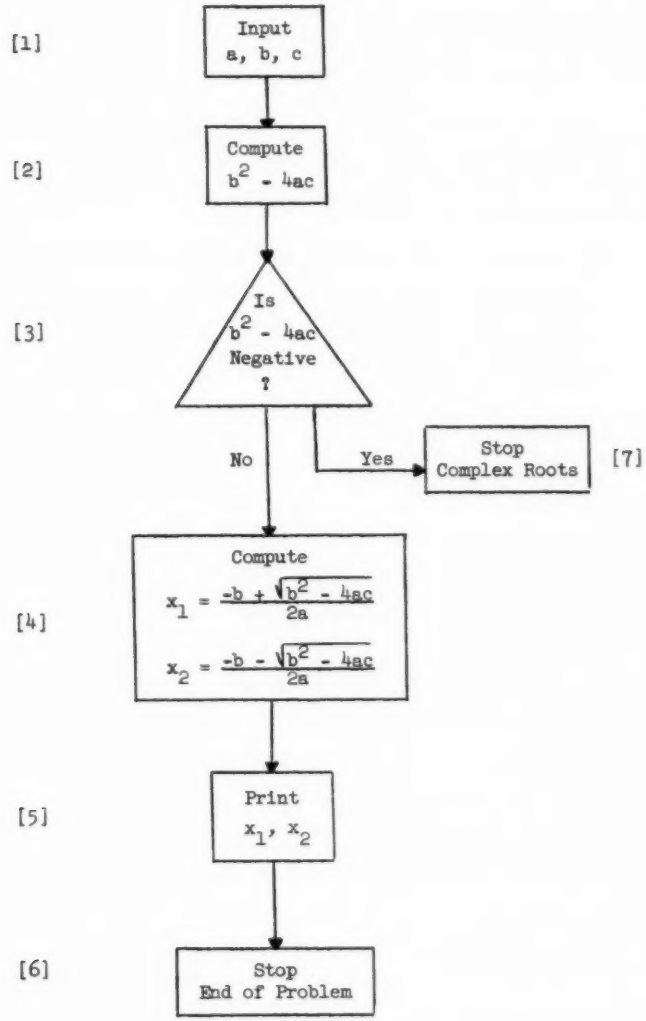
STEP III - PROGRAMMING

Programming is the development of the complete logical flow of the problem solution as follows:

Given: a, b, c

Find: x_1, x_2

Restriction: Interested only in real roots.


[1]
Input
a, b, c
[2]
Compute
$b^2 - 4ac$
[3]
Is
$b^2 - 4ac$
Negative
?
No
Yes
Stop
Complex Roots
[7]
[4]
Compute
$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$
$x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$
[5]
Print
x_1, x_2
[6]
Stop
End of Problem

STEP IV - CODING

It is now possible to write down, in the specialized language of the computer, the detailed sequence of instructions for solving the problem. We are given two sheets of paper on which the lines are numbered. Shown on the right are the Data Sheet and the Instruction Sheet.

Refer to the Data Sheet. At the beginning of a problem, numerical values of a, b, and c are recorded on lines 100-102. At the completion of the problem, the values of x_1 and x_2 will be recorded on lines 103 and 104. Intermediate results, as they are developed, will be recorded sequentially starting on line 105.

The columns on the Instruction Sheet are as follows:

Column (2) - Where (line number) to get the first number.
Column (4) - Where to get the second number.
Column (3) - What operation to perform on these numbers.
Column (5) - Where to put the result of this operation.
Column (6) - Reference information for the coder.

The numbers in brackets in the Explanation Column refer to the corresponding blocks in the Flow Diagram. Consider first Block [2]. The instruction on line 003 says, "Take a number from line 100 (a), multiply it by a number from line 102 (c), and store the result (ac) on line 105." The instructions on lines 004-006 are executed sequentially in a similar fashion to complete the computation of $(b^2 - 4ac)$.

The instruction on line 007 asks, "Is the number on line 108 negative?" If the answer is "no," go on to the next instruction in the sequence (008). If the answer is "yes," the next instruction to be executed is located on line 018.

The instructions on lines 008-014 are carried out in sequence to complete Block [4] of the Flow Diagram. This completes the arithmetic portion of the computation.

Input - Block [1]

Assume that our computer can read standard IBM punch cards. The instruction on line 000 says, "Read one card and record on line 100 the information punched in this card." Three cards are read in sequence in order to get the numerical values of a, b, and c into the computer.

Output - Block [5]

The instruction on line 015 says, "Print the number located on line 103 (x_1)."

The instruction on line 017 informs the computer that the problem has been completed.

The instructions numbered 000-018 and the three numerical constants on lines 020-022 completely describe this problem to the computer. These nineteen instructions and three constants constitute what is known as the computer PROGRAM.

DATA SHEET

(1) Line Number	(2) Contents	(3) Explanation
100	a	Input Data
101	b	
102	c	
103	x_1	Final Answers
104	x_2	
105	ac	Intermediate Results
106	4ac	
107	b^2	
108	$b^2 - 4ac$	
109	$\sqrt{b^2 - 4ac}$	
110	-b	
111	$-b + \sqrt{b^2 - 4ac}$	
112	$-b - \sqrt{b^2 - 4ac}$	
113	2a	
.	.	
.	.	
.	.	
020	4	Numerical Constants
021	0	
022	2	

INSTRUCTION SHEET

(1) Line Number	(2) First Number	(3) Operation	(4) Second Number	(5) Result	(6) Explanation	
000		READ		100	a into 100	[1]
001		READ		101	b into 101	
002		READ		102	c into 102	
003	100	MULT	102	105	ac	[2]
004	105	MULT	020	106	4ac	
005	101	MULT	101	107	b^2	
006	107	SUB	106	108	$b^2 - 4ac$	
007	108	NEGATIVE ?		018	If negative, go to 018	[3]
008	108	SQRT		109	$\sqrt{b^2 - 4ac}$	[4]
009	021	SUB	101	110	-b	
010	110	ADD	109	111	$-b + \sqrt{b^2 - 4ac}$	
011	110	SUB	109	112	$-b - \sqrt{b^2 - 4ac}$	
012	100	MULT	022	113	2a	
013	111	DIV	113	103	x_1	
014	112	DIV	113	104	x_2	
015	103	PRINT			Print x_1	[5]
016	104	PRINT			Print x_2	
017		STOP			End of Problem	[6]
018		STOP			Error	[7]

STEP V - CHECKOUT

Checkout is necessary in order to assure the correctness of the CODING. This is done by presenting to the computer a test problem for which we know the answers.

First, the program must be stored in the memory of the computer. One card is punched for each instruction. These cards contain the information from Columns (1)-(5) of the Instruction Sheet. There will be nineteen instruction cards. Similarly, the numerical constants from lines 020-022 are punched into three cards from Columns (1) and (2) of the Data Sheet. These 19 + 3 = 22 cards constitute the PROGRAM DECK.

If this program deck is inserted into the card reader of the computer and the LOAD button is pressed, the computer reads these cards in sequence and stores the information from each card into a memory location corresponding to the line number punched in the cards. Thus, the instruction from line 003 of the Instruction Sheet is stored into memory cell 003 in the computer, etc. The number associated with each memory cell is called its ADDRESS.

Three cards containing numerical values of a, b, and c are placed into the card reader. The computer is told that the first instruction is in memory cell 000 and the START button is pressed. The computer executes the first instruction which causes the contents of the first data card to be stored in memory cell 100. It then gets the next instruction from memory cell 001 and continues in sequence until the problem is completed.

Test Problem

$$a = 1, \quad b = 6, \quad c = -16$$

$$x_1 = 2, \quad x_2 = -8$$

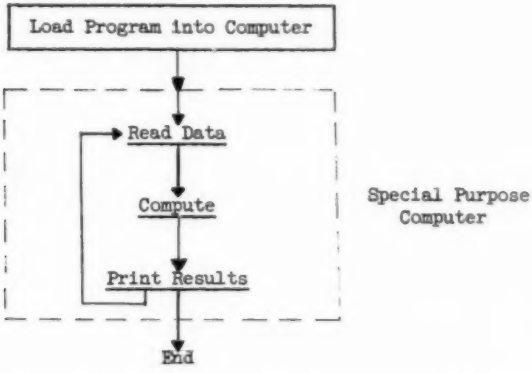
Corrections must be made in the program until the computer produces the correct results. Another test must also be made to verify that the computer operates correctly when $(b^2 - 4ac)$ is negative.

At this stage in a larger problem checks would also be made of the correctness of the Flow Diagram, the accuracy of any numerical approximations from STEP II, and the adequacy of the initial mathematical approximation to physical reality developed in STEP I.

STEP VI - PRODUCTION COMPUTATION

A digital computer is a general purpose device which is potentially capable of doing almost anything. Until it receives a set of instructions, however, it is actually capable of doing almost nothing. The process of reading a set of instructions into its memory converts a general purpose computer into a special purpose device which is immediately available to solve a particular problem—in this case a quadratic equation.

Below is shown the flow diagram of a typical production computation.



A computer would never be used to solve one quadratic at a time. It would be used if there were many sets of input data. The original program can be made to do this by replacing the STOP instruction (line 017) with the instruction: GO TO (000) which tells the computer to get its next instruction from line 000. This program will continue to solve quadratics until it runs out of data cards.

With the IBM 704 Computer, we can solve 100 quadratic equations in considerably less than one minute.

STEP VII - EVALUATION OF RESULTS

For the original problem:

$$a = 1, \quad b = n = 6 \text{ in.}, \quad c = -A = -16 \text{ sq. in.}$$

$$x_1 = 2, \quad x_1 + n = 8$$

$$x_2 = -8, \quad x_2 + n = -2$$

Both sets of results are mathematically correct solutions to the quadratic. Only the first is a reasonable answer to the physical problem.

The results of all machine computation must be viewed in the light of the physical situation which is being studied.

* * *

With this simple example, I have tried to demonstrate the important aspects of a digital computer. First, with a limited set of simple instructions, it is possible to solve large and complex problems—a large problem may require several thousand machine instructions. Second, the computer is able to choose between two courses of action where this choice may be based on the results of previous calculations. By compounding these two-way choices, it is possible for the computer to make complex decisions. A third factor is the computer's tremendous speed;

a 704, for example, carries out up to 40,000 instructions per second. The combination of these three factors gives the computer a high degree of flexibility. A digital computer is able to solve almost any problem which can be expressed mathematically. A digital computer is also extremely adaptable—it can be ready to solve a new problem in the length of time that it takes to read a new set of instructions into its memory from punched cards or magnetic tape.

COMPUTER APPLICATIONS

To show how a digital computer is used and to demonstrate the speed and flexibility, a few problems will be briefly described. There are two major classes of computer problems—those which can be solved by hand and those which cannot.

I. Problems which can be solved by hand.

A. Data Reduction - In the case of engine tests or experimental stress analysis, for example, simple calculations must be performed before the results of the test are usable. When done by hand, often only enough calculations are performed in order to plan the next set of tests. A computer on the other hand always carries out a complete set of calculations rapidly and accurately.

B. Component Design - Component design most often consists of carrying out calculations which evaluate the performance of a proposed component configuration.

In the case of a heat exchanger to be produced in large volumes, the objective is to design a unit which has satisfactory performance and which can be produced at a minimum cost. Both performance and production cost can be determined from the several independent variables. In one instance several thousand configurations were evaluated and the best combination selected. So many combinations were evaluated that the engineers were confident that they had arrived at an optimum design.

In the case of gear design, we now have a computer program which, without human intervention, modifies the geometry of the mating gears until an optimum is arrived at. A pair of matching gears can be designed in less than five minutes of computer time.

II. Problems which cannot be solved by hand.

In the problems just described, relatively simple computations were repeated many, many times. If done by hand, the job could be done more rapidly by carrying out several computations in parallel. In the analysis of complex engineering systems, however, such a large number of calculations must be performed sequentially that it would not be feasible to complete a single set of calculations by hand.

An example is the analysis of a vehicular gas turbine engine. An engine is made up of several interacting components. Even though engineers may be expert in the design of each of the components

which make up a gas turbine engine, it is not possible to predict how an engine will perform when a particular set of components are put together. The usual process is to build an engine, test it, modify it, retest it, etc., until a suitable combination is achieved. This is both expensive and time consuming. We have prepared a computer program which will completely determine the performance of a proposed engine in three minutes. The same calculations would take several months by hand.

In the development of the engine for the Firebird II, the computer was used to evaluate more than 100 proposed designs within the space of a few months. As a result, the development of a new engine was completed rapidly and the physical engine performed according to specifications. The use of a computer resulted in a large saving in elapsed time, and permitted engineers to learn rapidly about a new field.

THE ROLE OF THE COMPUTER IN STATISTICS AND QUALITY CONTROL

Quality control necessitates the gathering of vast amounts of data. Some of this data is usable in its raw form, but most of it must be processed in some manner. Because of the difficulties inherent in processing this data, much of the information contained in the data is never extracted and put to use. Much of it resides, untouched, in long rows of filing cabinets.

With a digital computer, it is now possible to subject this data to standard statistical tests. This can be done without a large expenditure of either time or money, since each computer installation has a number of flexible subroutines available which can be used by the statistician.

There are programs which calculate sample statistics such as mean, variance, standard deviation, correlation coefficient, higher moments, etc. There are routines which provide a least squares curve fit for a set of data points. There are routines for more complex analyses such as multiple correlation and regression, multivariate analysis, factor analysis, and linear programming. With these routines, a great deal can be learned from data which already exists. In the foundry, for example, studies can be made to determine which are the important variables, and of equal interest, which are the unimportant variables, and what is the relationship between these variables and the end product.

One deterrent to the immediate use of a computer is the fact that source data is not in a form suitable for mechanical processing. Future tests should be planned with a computer in mind. Data should be taken in a form for easy conversion to punch cards so that all further processing on this data can be carried out mechanically.

The digital computer is perhaps the most useful tool available to the statistician and quality control analyst. It is a tool which many of you can use now. There are several problem areas in which a computer becomes a necessity.

One such area is the design and manufacture of complex systems. Like the gas turbine engine whose performance depends on its interacting components, so control limits can be adequately specified only if their

affect on the over-all system is known. The affect of component variation can be tested in advance using a mathematical model of the complete system.

Machine tools are now available which operate automatically under the control of magnetic tape. A digital computer is used to prepare this control tape from dimensions on drawings. When these machines are used for volume production, it will be necessary to build quality control into the computer program.

I am sure that you will find your future with computers both interesting and rewarding.

QUALITY CONTROL AND PRODUCTION EFFICIENCY

Ralph D. Humphries
Cessna Aircraft Company

An outstanding discovery has been made in modern industry. Business and industry alike are turning to quality control for production efficiency. The results of quality control and production efficiency are synonomous.

The object of any company is to produce a product or a service, to supply a ready demand or a created demand. It must have value and/or customer appeal. The characteristics associated with the product or service must possess quality. Quality is a broad term. We normally think of the quality of a completed product, but in order to develop, produce and deliver successfully, we must maintain some form of quality control of all operations. This is a relatively new concept. A few years ago production efficiency was considered to be a high rate of productivity at a low cost. A high rate and low cost is still becoming more important - but another very influencing variable has entered the picture, "Quality."

Production efficiency can not only be measured by the speed and cost of productivity. The production results must be measured. Speed, cost and quality of the product must be properly balanced. To balance these three factors, a different quality must be considered - the quality of factors responsible for the cost, the productive speed and the quality of the product. The many detail thoughts and actions, responsible for the creation of a product, must possess a degree of quality.

There are two important aspects of the term Quality Control.

1. If we use the term as a noun, we are referring to a department or a method. To me, this seems to be the most common usage. The department is responsible for seeing that a satisfactory product is delivered to the customer. Frequently, the department has little or no control over how the product is to be produced. Their function is to separate the good from the bad. If the pile of bad is too big, it is time for a review of the specifications or the method of inspection.

2. If we use the term as a verb, it becomes a part of any act or operation. All phases of operations require some standard of quality. Engineering, Planning, Tooling, Manufacturing, Inspection, Accounting, Sales and all related detail paper work must possess quality. The control of such quality must be planned. It must be premeditated or become a basic philosophy of everyone involved.

Both aspects of quality control are normally required; however, the more completely the planning and performance are controlled, the more inspection can be reduced.

Quality must be a basic ingredient built into a product or a service. It is not a frill or a mere surface characteristic. We cannot just spray paint it on or inspect it into a product or service. Quality is a basic ingredient for which the customer pays, expects, and deserves.

Employees should understand which quality standards they are expected to meet and why close adherence to quality standards is so important in maintaining customer acceptance and good will.

The concept of quality control is largely determined by the management attitude. The employees are motivated to a great extent by the environment, the indoctrination, and individual supervision within the organization. Management and employees who have the quality-mindedness need a guiding hand. They need signs to show the way.

We at Cessna feel that the communication between inspection, management and the quality producing departments or employees is very important. The communication systems and procedures must be simple, economical, and efficient.

To illustrate how quality control can influence production efficiency, let us review a few examples.

In Figure 1, the shop routing card reflects the emphasis placed on the crew chief's responsibilities. Each crew chief is assigned a stamp indicating the department, shift and section number. Each operation is stamped off before it is moved to the next section.

SR 1990

ENG. CHANGES	SHT	SERIALS	NEXT ASSEMBLY	N. A. NAME				ITEM NUMBER	QUANTITY
NC		35236 & ON	0813300-90	U Lock Assy	1	1	1		

ORDER number C-8 DUE DATE

PART NUMBER 0813301-1

PART NAME Rod End Assy

SECTION Nose MODEL 310

SUPERSEDES 0813300-85

MATERIAL DESCRIPTION	MATERIAL CODE	STOCKROOM	QTY. PER PART	U. OF M.	TOTAL MATERIAL
D (1) 0813301-2 Rod End (MOA) 1-1/16 Dia. 4130 Steel Rod	8141915	57-2	1.13	Lbs.	
P (1) NAS508-4 Bearing	J034230	57-1	1.00	Ea.	

APPROV. PART DIM. | PROGRESSION | PARTS PER STOCK | STRIPS PER SHEET | PARTS PER SHEET | ROUTING REVISION DATE JM1-17-57 | ROCKWELL C 86 | BARON | WEBSTER

MANUFACTURING PLANNER

OPER	DEPT	SECTION	SHIFT	STANDARD HOURS SET UP	STANDARD HOURS 100 PCS	TOOLING	MANUFACTURING OPERATIONS
10	22	06	01	1.15	6.882	SW 129	Turn O.D. .500 and Chamfer .03 x 45° and Thread 1/2-20 UNF-3A x 2.70. Cut to length 4.35
20	22	06	01	.58	5.430	FT 20089 FT 20052	Form .50 Radius and .500 Sphere Radius to Net Length 4.20
30	64						Inspect
40	34						Heat Treat to R.C. 34/40
50	34						Sand Blast
60	64						Inspect
70	22					SW 129	Chase Threads 1/2-20 UNF-3A x 2.70
80	22	07	51	.74	4.923		Mill Sphere .505 Thick Feed 1" RPM 41
90	22	04	03	.27	2.260	DJ 72375	Drill .609 Hole
100	22						Broach .8245 +.0010 -.0005
110	64						Inspect
120	34						Cad Plate Per CES-1081
130	34						Bake
140	64						Inspect
150	64						Magnaflux
160	22						Press in Bearing and Stake 6 Places
170	64						Inspect
180	22						Count Point
▸190	52	06					Store

Figure 1

In Figure 2, we have developed an inspection procedure card used by the detail inspector as a guide for sampling and inspection instruction. It is also used as a file record for all orders received. This form is versatile so that it may be used in all departments. Accumulated information can be summarized on the card at any time. However, we have found that an easier way to summarize the results is through I.B.M. tabulation.

FORM 874 — INSPECTION PROCEDURE CARD

PART NO 6-35016-2	PART NAME LOCK ARM						CONTRACT B-47		CODE N603456			
SOURCE 2025	DATE	6-8-56	6-26-56	8-30-56								
PLANNED BY / INSP. SUPERVISOR	ORDER NO.	7-28163	7-33001	7-39455								
CUSTOMER APPROV.	R. R.	442/3	5T429	55317								
	LOT SIZE	60	435	200								
LATEST CHG. "B"	TABLE	1	2	3	4	5	6	7	8	9	10	
Test reports on file at vendors	R/R	8-0	14-0	10-0								
All processing done by gov't approved sources	R/R	8-0	14-0	10-0								
Vendor acceptance stamp	12	8-0	14-0	10-0								
Steel stamp part number	12											
Code number	12											
Cadmium plate per flag (3) (visual)	12											
Involute spline checked per D-7539	12											
57° ±1/4° angle	12											
2.29/2.30 dim.	12		#2-0									
.250/.254 dim.	12		14-0	10-0								
Scribe line (visual)	12											
.125 dim.	12											
Rockwell	12	8-0	14-0	10-0								
	CODE											

Figure 2 (front)

REMARKS

LOT NO.

1 O.K. ✓

2 2.29/2.30 O.D. IS U.S. .001/.0015 ✓

3 .250" DIA HOLE O.S. .0012 ✓

4

5

6

7

8

9

10

Figure 2 (back)

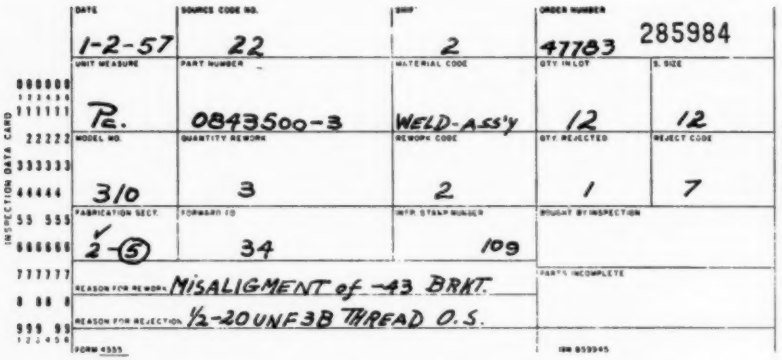


Figure 3

Figure 3, the I.B.M. card, is the most versatile form used. For each order or lot of parts inspected, a card is filled out with the information necessary to tabulate a summary as required. This card is developed for use in any department of production. Code numbers are used to identify vendors, shop discrepancies, materials, etc. In developing a code, it is advisable to coordinate with all other departments which may use the same code. For example, the vendor or source code is used by Purchasing, Accounting, and Inspection. The rejection code

should be one which is easy to use and to interpret. In a quality minded plant, the supervision usually is aware of the errors. They may employ various ways of keeping informed of the mistakes being made. In our shops, the foreman who is responsible for all production in each department requests that he be allowed to sign all rejection tags and rework orders. Now, due to the jumble and variety of discrepancies throughout the day, he needs a tally or recap over a period of time to determine how many like parts with like discrepancies are being rejected and in what area the greatest degree of corrective action is needed. It is for this purpose that a code is developed - one that reduces numerous detail characteristics into a few broad-term characteristics. For example, see Code Designations.

Rejection and Rework Code Numbers

1. Wrong Raw Material: type or thickness; tensile strength and chemical analysis unsatisfactory; no certification; certification not available; blisters; flaws.
2. Dimensional Error: too large or short; too wide or narrow; too thick or thin; too deep or shallow; oversize or undersize.
3. Forming Error: misformed flanges; formed backwards; joggled wrong; dimpled wrong.
4. Drilling and Reaming Error: out of line; misaligned; oversize; undersize; rough; elongated; tapered holes omitted; double drilled; drill runs; drill marks; drilled in error.
5. Countersinks and Spot Faces: too deep or shallow; elongated. rough surface; cocked; countersunk or spot faced in error.
6. Machining Error: undercut; cuts not parallel; off in concentricity; not perpendicular; finish rough or wavy.
7. Threads: wrong type; rough; out tolerance; not concentric; tapered; threaded in error.
8. Welding and Brazing: wrong type weld or braze; operation in error; insufficient penetration; pin holes and flaws; undercuts and burnt material.
9. Protective Treatment: defective plating; anodize; iridite; paint; primer; unprotected; operation in error.
10. Corroded: pitted; porous; cavities; rusted; stained.
11. Operations: missed; reversed; wrong; incomplete.
12. Trimmed; rough trim; cutouts in wrong location; undersize and oversize; trimmed in error.
13. Oil Canned: buckled; bowed; twisted.
14. Material Condition: scratches; rough or wavy surface; torn; dented; cracked; scribe marks; handling damage; blisters; delaminations; won't harden; stains; proof load low.
15. Heat Treat and Age Harden: operation incomplete; not within specification; heat treated too many times; no test coupons; operation in error.
16. Contour: wrong contour; under; over; wavy.
17. Identification: wrong number; number poorly applied; wrong method of identification; identification obliterated; damaged by identification.
18. Rivets: bad; missing; wrong pattern; wrong size or type; too many; incomplete; not in shear.
19. Assembled: wrong per specification; parts missing; improperly located; out line; with wrong bolts, screws, nuts, and bearings; defective staking; omitted parts; failed functional test.
20. Riveting Damage: set marks; bar marks; dings; gouges.
21. Bad Repair: unauthorized repair; cannot be or was not repaired per disposition.
22. Electrical Error: damaged wire instruments and units; will not operate; burnt out in test; not per specification; not properly insulated and installed.
23. Misfits: tight assembly; overlaps of tolerances; gaps; misaligned; bad appearance; ruined in trying to assemble.
24. Planning Operation: left out operation; in wrong sequence; different than specification; not made for correct serial; too many ordered.
25. Engineering Error: drafting error; cannot be assembled; will not function; failed under load; cannot be made per specification; redesigned.
26. Tooling Error: inadequate tools; no tools available; tools in error; parts made by tooling department; tools not per specification.
27. Purchasing: purchased wrong parts; error on purchase order; purchased from unapproved vendor.
28. Inspection: bought parts in error; ruined the parts by inspection.
29. Handling: delivered to wrong place; bypassed departments; damaged in transit; not properly protected in transit; improperly stacked.

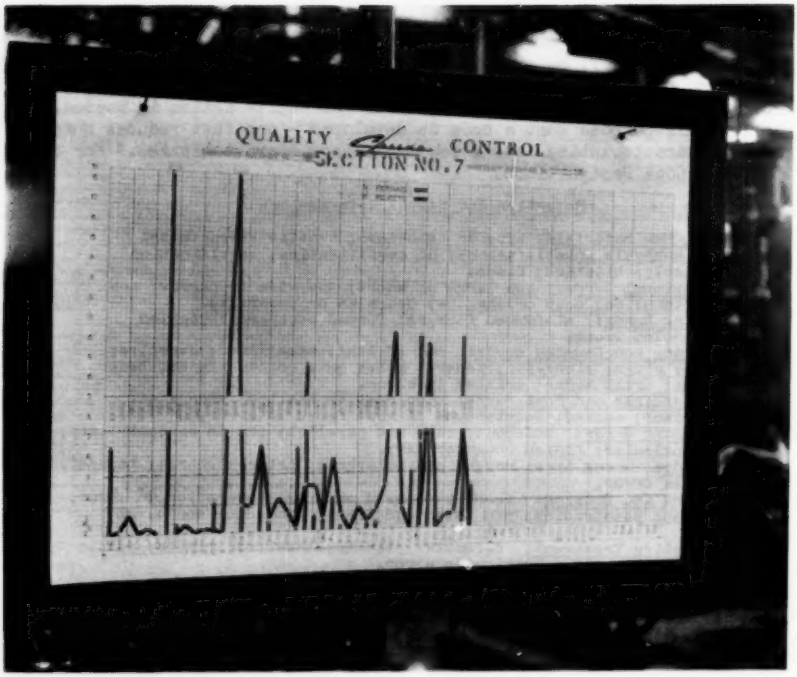


Figure 4

To keep the productive shop informed on the quality efficiency, a chart is located in each department section. A section is a designated area supervised by a crew chief. The chart is plotted daily from the results recorded on the I.B.M. cards. The chart reflects the percentage of rework and rejection found by inspection. See Figure 4.

To encourage corrections by the production departments, we developed a shop scrap tag. This tag is used only by the shop personnel to identify material which they know is defective and do not wish to present to inspection. The items they reject are not entered in the tabulated report or charted with the other inspection rejections. They are accumulated with other scrap items in Salvage. In this way a quality incentive is created - a desire to check their work more closely.

Once a month all cards from each department are tabulated into two reports.

One tabulation shows the total parts presented to inspection, total parts rejected, quantity of lots presented and quantity of lots rejected on each part number in a department or from a vendor. This is accumulative month to month for six months.

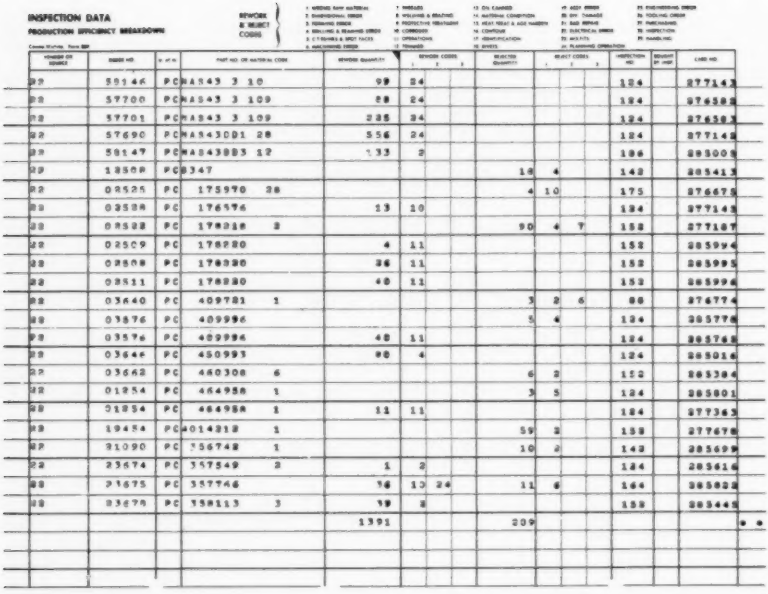
INSPECTION DATA

ACCUMULATED REPORT

	Period 1st thru 3rd				
	October 31st				
	1955				
22	0300106-30	25		1	
22	0300106-31	25		1	
22	0310110	12	12	1	1
22	0310134-1	252	252	2	2
22	0310146	116	34	1	1
22	0310188-4	7		1	
22	0310295	1266		2	
22	0310296	1146		4	
22	0310419	10		1	
22	0341109-1	220	4	11	2
22	0411579-5	469		7	
22	0842000-12	305		8	
22	0842000-13	153		4	
22	0842000-30	339	17	10	2
22	0842000-17	1003		11	
22	0852250-15	363		2	
22	0861100-5	180		3	
22	0861300-6	72	2	2	1
22	1050107	111		2	
22	1051000-10	45		1	
		705320	11239	6703	511

Figure 5

A second tabulation shows disposition of all defective lots, listing for each department or vendor the date, quantity rejected, reason for rejection, quantity returned for rework and reason for rework for each part number. See Figure 6.



INSPECTION DATA

PRODUCTION EFFICIENCY BREAKDOWN

REWORK & REJECT CODES

VENDOR OR SOURCE	INDEX NO.	U. of M.	PART NO. OR MATERIAL CODE	REWORK QUANTITY	REWORK CODES 1	2	3	REJECTION QUANTITY	REJECT CODES 1	2	3	INSPECTION NO.	REQUEST BY INSP.	LOSS NO.
22	58146	PC	NAS43 3 10	99	24							124		277143
22	57700	PC	NAS43 3 109	28	24							124		276582
22	57701	PC	NAS43 3 109	225	24							124		276593
22	57690	PC	NAS43001 28	556	24							124		277142
22	58147	PC	NAS43003 12	133	2							126		285005
22	18568	PC	8347					18	4			142		285413
22	02525	PC	175970 28					4	10			175		276675
22	03538	PC	176576	13	10							124		277145
22	03522	PC	178218 2					90	4	7		152		277187
22	02509	PC	178220	4	11							152		285994
22	02508	PC	178220	36	11							152		285995
22	02511	PC	178220	40	11							152		285996
22	03640	PC	409781 1					3	2	6		98		276774
22	03576	PC	409996					5	4			124		285776
22	03576	PC	409996	48	11							124		285768
22	03646	PC	450993	98	4							124		285016
22	03642	PC	460308 6					6	2			152		285384
22	01254	PC	464958 1					3	5			124		285001
22	01254	PC	464958 1	11	11							124		277363
22	19454	PC	4014218 1					59	2			138		277676
22	31090	PC	556748 1					10	2			142		285699
22	23674	PC	557549 2	1	2							124		285614
22	21675	PC	557746	16	13	24		11	6			164		285822
22	03679	PC	558113 3	19	2							152		285445
				1391				209						

Figure 6

Charts are sent to supervision and management monthly showing the trend in rework, rejection, scrap and quality efficiency. See Figures 7, 8, and 9.

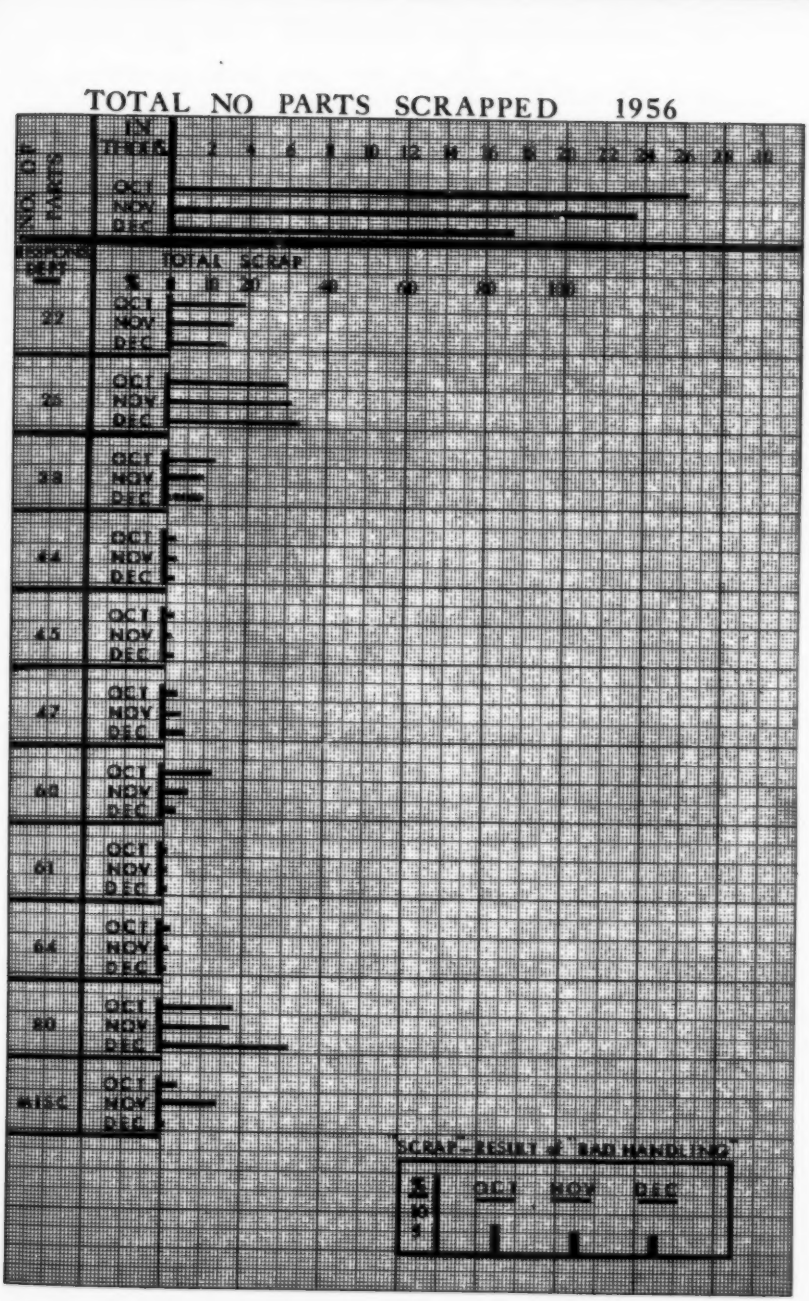

TOTAL NO PARTS SCRAPPED 1956
NO. OF PARTS
IN THOUSANDS
OCT
NOV
DEC
DEPT. NO.
TOTAL SCRAP
$
27
26
38
44
45
47
48
61
64
80
"SCRAP" RESULT OF "BAD HANDLING"
$
OCT
NOV
DEC

Figure 7

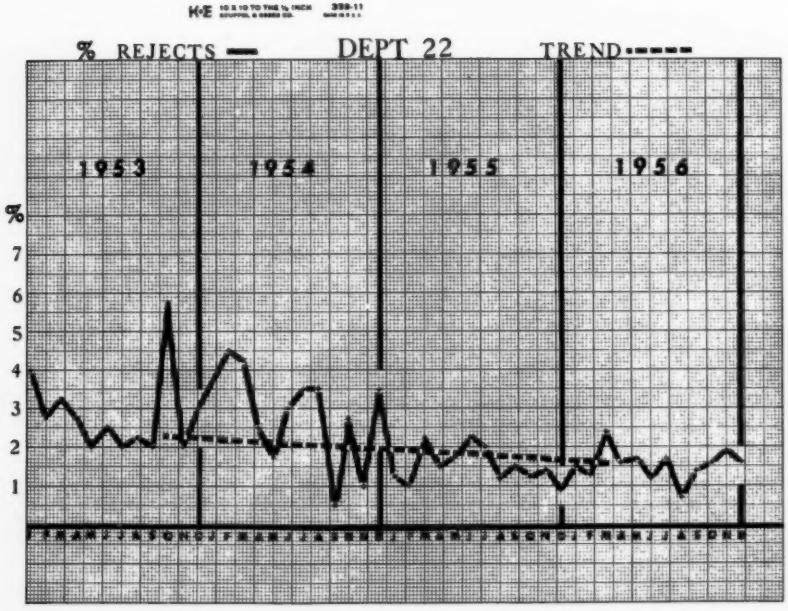


Figure 8

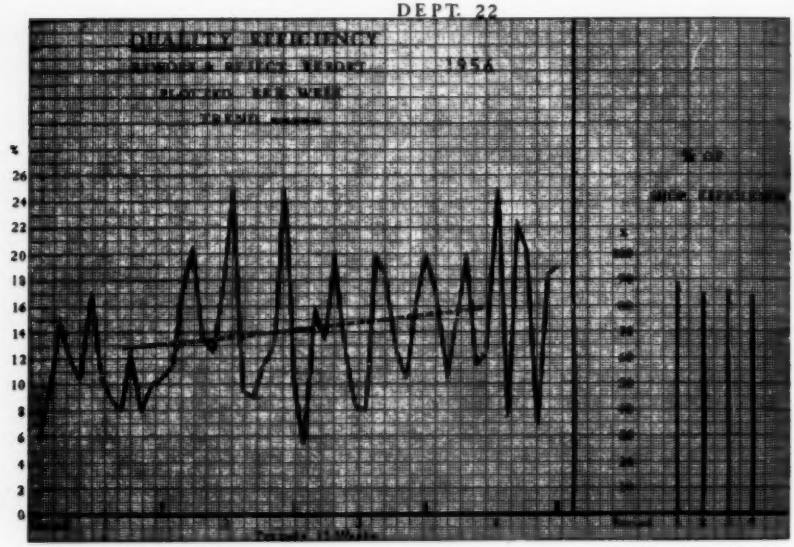


Figure 9

The effectiveness of the system used and how much improvement is required is also governed by applying the same method to the finished and delivered product.

In Figure 10 the chart in final assembly reflects the condition reported by final inspection. The total number of squawks is recorded for each airplane. The chart is located on the side of the foreman's office where everyone can see the department's performance record. The inspectors fill out an I.B.M. card shown in Figure 3 on all rejected parts. The information is analyzed to determine the cause for rejection and to make the necessary corrective action.

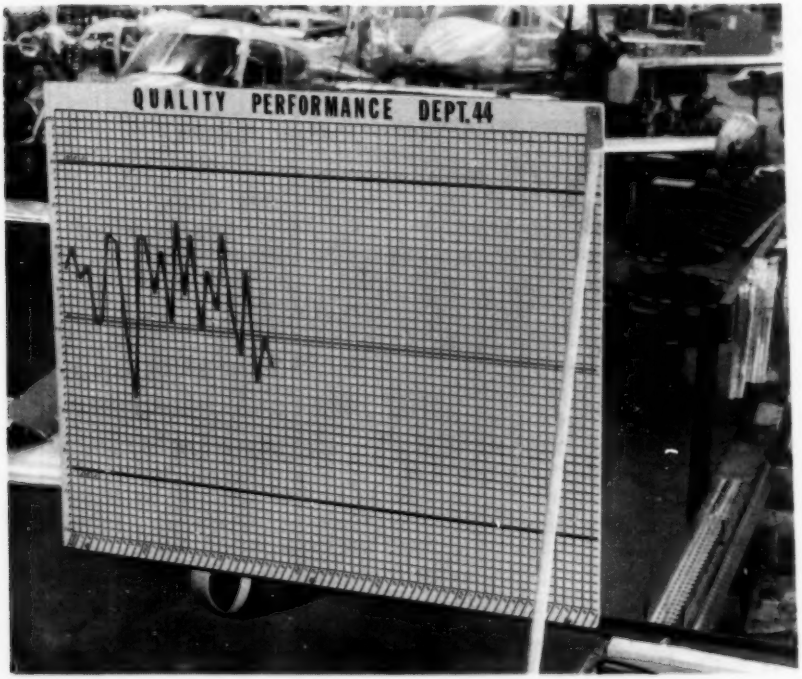


Figure 10

The final acceptance of the airplane is made by the test pilot. The pilots reports on all airplanes are summarized monthly to determine the area and importance of the re-occurring squawks. In Figure 11, the chart indicates in what area the most emphasis on corrective action should be placed.

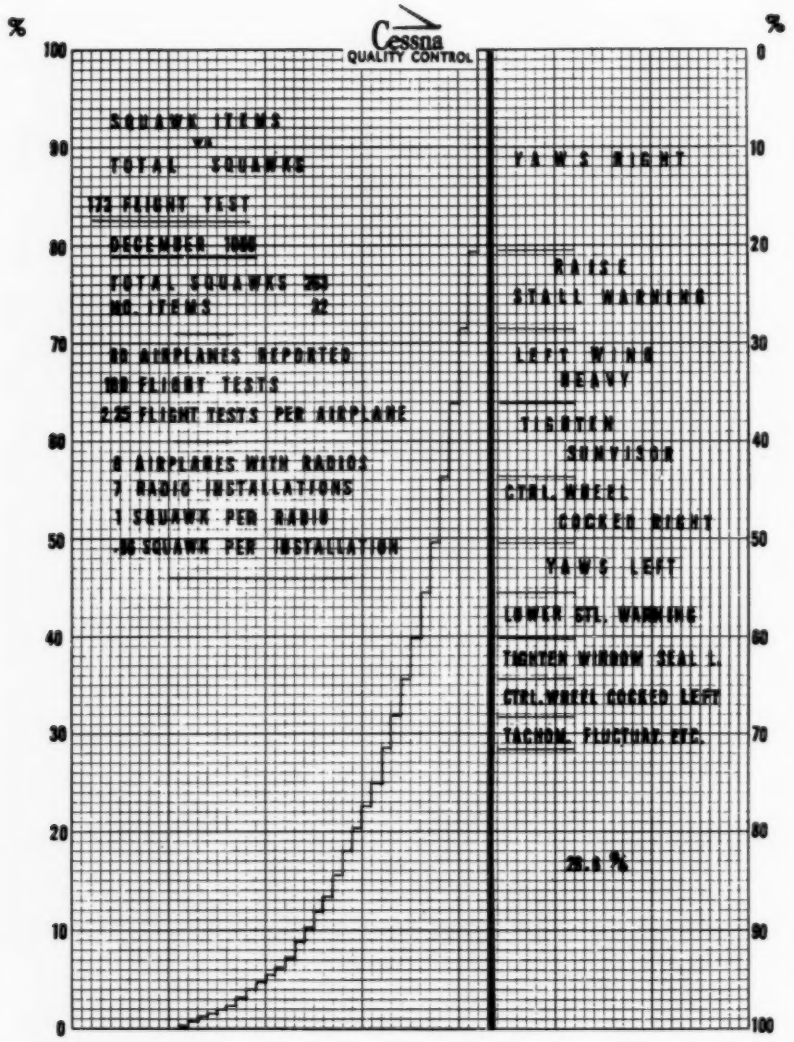


Figure 11

Customer reports are summarized to determine the characteristics found unsatisfactory after delivery. In Figure 12, the chart shows the items found most frequently in the reports. By comparing the charts, a much more realistic decision can be made on our quality index.

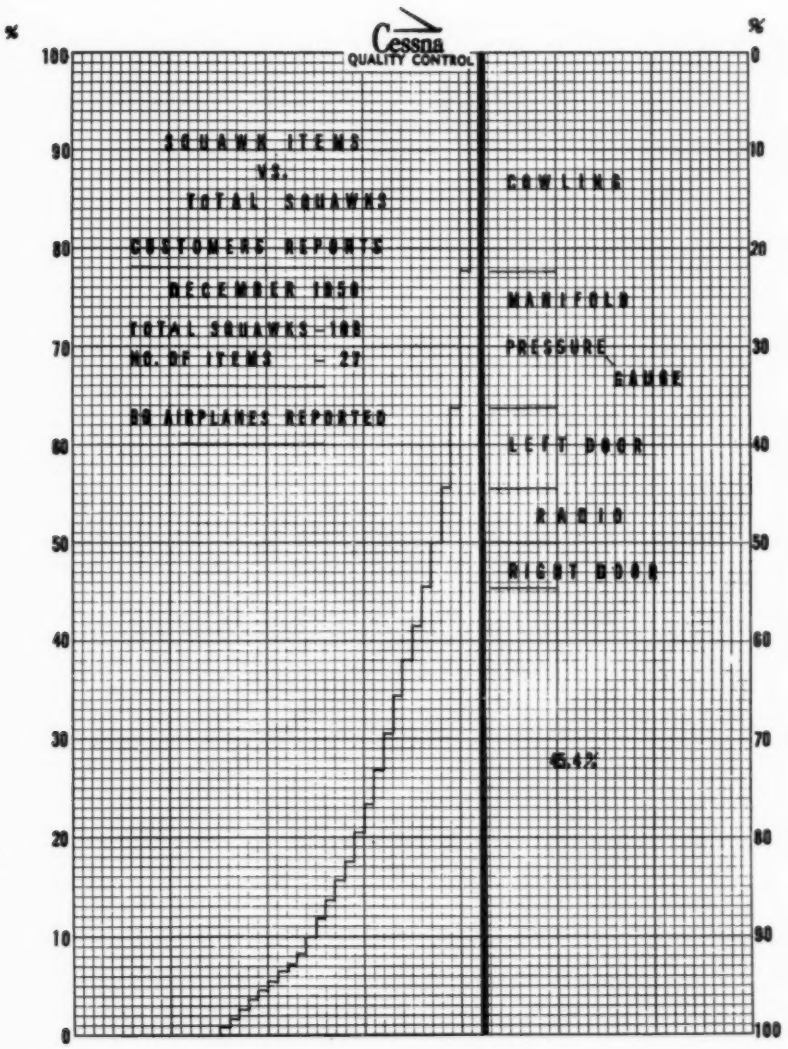


Figure 12

At this point, we have developed a chain of communication between the product producing departments and the consumer.

We see what we have done.
We see where we are going.
We see where we can make our best improvements.

For production efficiency, we must control the most important factor we have to sell, a "quality product" - one that creates a desire to own, to admire and to use.

If we have the desired quality in all production operations, we have the desired production efficiency.

SOME USEFUL NONPARAMETRIC SIGNIFICANCE TESTS

Fred A. Beeler
Western Michigan University

It is the aim of this paper to discuss three nonparametric significance tests and show how they can be of use to personnel in the field of Quality Control.

In many of the tests of significance used in the field of statistics we assume that the population has some definite form. Frequently we assume the normal distribution and we estimate or test hypotheses about the mean and variance. For example when use is made of the Student-Fisher t it is assumed that the sample is random and is drawn from a normal universe. The mean and standard deviation of a normal population are called parameters of the distribution. The estimation or testing hypotheses concerning these parameters of a population are called parametric statistics.

In the last few years advances have been made in finding new test statistics which will compare two distributions without specifying the form of the distribution. These methods are called nonparametric statistics or distribution free statistics. The only assumption that is needed for most of the nonparametric tests is that the frequency function be continuous. The best known and probably the most useful of the nonparametric techniques are the χ^2 and the binomial tests. Since these are well known no more mention will be made of them here.

Before we take up the examples illustrating the use of some nonparametric methods it seems wise to reconsider the idea of measurement and also to define what is meant by the efficiency of a significance test.

Measurement. There are four categories of measurement: they are the nominal (classificatory), ordinal (ranking), interval, and ratio scales (1).

1. Nominal or classificatory scale. This is measurement at its weakest level of existence when numbers or other symbols are used simply to classify an object, person or characteristic.

   Examples: Classifying paper by its color.
   Classifying automobiles by makes.

2. Ordinal or ranking scale. It may be that the objects of a category in the nominal scale stand in some kind of relationship that may be designated by the carat (>) which means greater than, is preferred to, is higher than, is more difficult than, etc.

   Examples: Sergeant > corporal > private.
   Scores obtained on aptitude and ability tests.*

*(Very few ability test scores would warrant a more precise scale of measurement. Who can justify that a score of 80 is twice as good as 40? Who can justify that the increase in ability to increase a score from 40 to 50 is the same as the increase from 90 to 100?

3. Interval scale. When a scale has all the characteristics of an ordinal scale and in addition the DISTANCE between any two numbers on the scale are of known size, then the measurement is considerably stronger than that of the ordinal scale. In this sort of measurement, the ratio of any two intervals is independent of the unit of measurement and the zero point.

   Examples: Temperature (Fahrenheit and Centigrade)
   Gauge pressure
   Excess above 3.450 in. in the diameter of a bearing, given in .001 in.

4. Ratio Scale. When an interval scale has a true zero then we have a ratio scale.

   Examples: Temperature (absolute)
   Distances (starting from zero)
   Weight
   Cost of an item.

In the fields of physical science the interval and ratio scales are usually used. In the behavioral sciences the nominal and ordinal scales are usually used.

In the selection of a significance test a person must select a test that is appropriate to the level of measurement. For data measured in the nominal and ordinal scales one usually uses a nonparametric test as there are few parametric tests available. For data measured in the interval or ratio scales the parametric methods should be used if the assumptions of the parametric statistical model are tenable. If one uses a nonparametric test when a parametric test could be applied, then the parametric test is usually more efficient. For example, if a nonparametric test is 80% efficient, one means that if a sample of 100 is used in the nonparametric test to reach a decision then a sample of only 80 would be needed by the parametric test to reach a decision.

In the past few years a number of nonparametric tests have been developed but unfortunately they have been published in various journals and sometimes in a highly technical form making them almost inaccessible to many people working in Quality Control. This disadvantage has to a large degree been overcome by a book published by the McGraw-Hill Book Company this last fall. It is a textbook by Sidney Siegel called "Nonparametric Statistics for the Behavioral Sciences." While this book was written for people working in the behavioral sciences, it is the first book on the subject and the author does a creditable job in presenting the various tests in non-mathematical language. There are tables in the back of the book to go with each test. Also some of the newer textbooks devote some space to nonparametric statistics. Some of these books are listed in the references.

The three nonparametric tests that will be discussed are:

1. The sign test,
2. The Wilcoxon matched-pairs signed-rank test,
3. The corner test of association.

THE SIGN TEST

To explain this test let us consider the data given in table I. Here we have two determinations of the chlorine demand on each of 11 lots of paper pulp, one determination is made by the vandor and the other by the consumer.

TABLE I

CHLORINE DEMAND MADE BY VENDOR AND CONSUMER
(standard test)

Lot	Vendor analysis	Consumers analysis	$d = X_V - X_C$
	V	C	
1	3.5	3.7	-
2	5.5	4.1	+
3	5.4	3.8	+
4	4.6	4.1	+
5	4.9	3.8	+
6	5.6	4.0	+
7	4.0	3.8	+
8	4.9	4.2	+
9	4.9	3.9	+
10	4.5	4.2	+
11	3.6	3.6	0

The hypothesis that we want to test now is that the median of the differences of the analyses is zero. To make this test determine the sign of $d = X_V - X_C$, i.e. from the vendors determination subtract the consumers determination and record only whether this difference is positive or negative. These signs are given in the last column of table I. Any difference that is zero is deleted from the test.

When the hypothesis that the median of the differences is zero is true then one would expect about as many positive differences as negative. Whenever there are too many positive (negative) differences one doubts that the hypothesis is true.

In our example there are 10 differences that are not zero so take N=10. There are 9 plus and 1 negative signs. The probability of obtaining 9 or more plus signs or 9 or more negative signs is 0.022. Tables giving the probabilities for this test can be found in (1,2,3). This is exactly the same problem as tossing 10 coins and asking what is the probability of getting 9 or more heads or 9 or more tails in a single toss. At the 5% level of significance we will reject the hypothesis; that we will conclude that the vendors and consumers determinations do not agree. However at the 1% level of significance one must accept the hypothesis since we could not reject it.

The efficiency of the sign test (2) when compared to two normal distributions with a common variance is about 96% for samples of 6 and decreases to about 70% for samples of size 20 and gradually decreases to about 63% for very large samples.

In this example we call the two chlorine analyses matched samples since they are made on the same lot of pulp. There is variation in the chlorine demand for the different lots but by making two determinations

on the same lot and subtracting them when this variation between lots does not appear. All that is needed in the way of measurement for this test is a partially ordered ordinal scale.

This test is most useful when we have

1. Measurements on two things to be compared,
2. Each of the measurements of the pair are made under similar conditions,
3. The different pairs were made under different conditions.

WILCOXON MATCHED-PAIRS SIGNED-RANKS TEST

This test is similar to the sign test except that now we will make use of the magnitude of the differences obtained from the matched pairs. Of course in this case one needs a more refined category of measurement. This Wilcoxon test requires ordinal measurements not only within pairs as required for the sign test but also the differences between pairs. While this test is more efficient than the sign test it also requires more in the way of measurement. This Wilcoxon matched-pairs signed-ranks test is about 95% efficient as compared to the Student-Fisher t test. Of course this implies that the category of measurement is at least an interval scale and that all of the assumptions about normality must be satisfied in order to compare this test to a t test. If our category of measurement is only in the ordinal scale then the t test cannot be used as it does not exist for this type of measurement.

To illustrate the details of performing this test let us again consider the same data used for the sign test.

TABLE II

CHLORINE DEMAND MADE BY VENDOR AND CONSUMER
(standard test)

Lot	Vendors analysis V	Consumers analysis C	$d=X_V-X_C$	Rank $\|d\|$	Positive Signed rank	Negative Signed rank
1	3.5	3.7	-.2	2.5		-2.5
2	5.5	4.1	1.4	9	9	
3	5.4	3.8	1.6	10.5	10.5	
4	4.6	4.1	.5	5	5	
5	4.9	3.8	1.1	8	8	
6	5.6	4.0	1.6	10.5	10.5	
7	4.0	3.8	.2	2.5	2.5	
8	4.9	4.2	.7	6	6	
9	4.9	3.9	1.0	7	7	
10	4.5	4.2	.3	4	4	
11	3.6	3.6	0	1	1	
				Total	63.5	-2.5

To make this test do the following steps illustrated in table II.

1. Obtain the difference $d=X_V-X_C$.
2. Rank the absolute values of the d's. (in case of a tie, record the average of the ties).
3. Place a negative sign on the rank of $|d|$ if d is negative.
4. Sum the positive and negative ranks separately.

5. The test statistic T is the absolute value of the smaller sum of the positive and negative ranks. T = 2.5 in our example.

The hypothesis that one wants to test is that the chlorine demand as determined by the vendor and consumer are the same. Or to state it in another way, it is that the median of the differences is zero. If this hypothesis is true then one would find about the same number of positive differences as negative. So far this is the same as the sign test. In order to take into account the magnitudes of these differences, they were ranked in absolute value. Now if the hypothesis is true, one would expect the sum of the positive and negative ranks to be about equal. Since the total sum of ranks is fixed for any problem knowing the sum of positive ranks fixes the sum of the negative ranks. So the smaller sum has been taken as the test statistics. Since there are 11 pairs, N = 11. Looking in the table (1,3) of critical values of T, one finds at the 1% level of significance T = 5. Since our T is smaller than 5 we will reject the hypothesis. This means that one would say that the analysis by the vendor and consumer do not agree.

It should be noted that both the sign test and the Wilcoxon matched-pairs ranked-sign test are very easy and quickly done. There is no assumption as to normality nor is it assumed that they have a common variance even if normality exists, as is necessary for a t test.

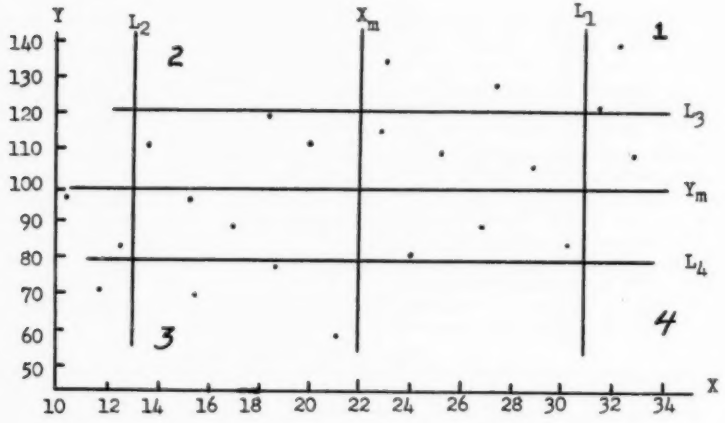
THE CORNER TEST OF ASSOCIATION

Suppose that N (even number) observations of a bivariate distribution are obtained and one would like to know if the two variates are independently distributed. For this test we assume continuity and we also assume that our measurements are at least in the ordinal scale. The corner test of association is easy to perform and will be explained by an example. Consider the X and Y measurements in table III.

TABLE III

Rank of X	X	Y	Rank of Y
1	10.2	97	12
2	11.6	71	20
3	12.4	83	17
4	13.6	111	8
5	15.2	96	13
6	15.4	69	21
7	17.0	88	15
8	18.4	120	5
9	18.6	78	19
10	20.0	112	7
11	21.0	58	22
12	22.8	115	6
13	23.0	135	2
14	24.0	80	18
15	25.2	109	9
16	26.8	89	14
17	27.4	128	3
18	28.8	105	11
19	30.2	84	16
20	31.4	122	4
21	32.2	139	1
22	32.8	108	10

Scatter diagram for the X and Y measurements of table III



To make this test, do the following:

1. Rank the X measurements to obtain the median X_m.
2. Rank the Y measurements to obtain the median Y_m.
3. Plot the points on a scatter diagram.
4. Draw a vertical line through X_m.
5. Draw a horizontal line through Y_m.
   (There will be 11 points on each side of the vertical line through the median of X and also there will be 11 points above and below the horizontal line through the median of Y. These two lines will divide the points into four quadrants which will be numbered I through IV in the usual way. In making this test the points in quadrants I and III will be taken as positive and the points in quadrants II and IV will be taken as negative.)
6. Take any vertical line which is to the right of all points and move it to the left until a point is reached which is on the opposite side of the horizontal median line from the furthest point on the right. Call this line L_1. Count the number of points to the right of L_1 and denote the number by r_1. If these points are in the first quadrant r_1 will be taken as positive and if they are in the fourth quadrant r_1 will be taken as negative. In our examrle $r_1 = 3$.
7. In a similar way bring a vertical line in from the left and call this line L_2. Count the number of points to the left of L_2 and denote the number by r_2. This number will be taken as positive if the points lie in the third quadrant, otherwise it will be taken as negative.
8. In a similar way bring a horizontal line down from the top and another up from the bottom calling these lines L_3 and L_4. Count the number above L_3 and denoting the number by r_3 which will be taken as positive if the points lie in the first quadrant, otherwise it will be taken as negative. Also count the number of points below L_4 and denoting the number by r_4 which will be taken as positive if the points lie in the third quadrant.
9. Compute the test statistic $R = r_1 + r_2 + r_3 + r_4 = 3 + 3 + 4 + 4 = 14$

10. No tables are needed for this test for
    a) if $N \geq 10$ (22 in our example) and $R \geq 11$ or $R \leq -11$ we will reject the hypothesis of independence at the 5% level of significance,
    b) if $N \geq 10$ and $R \geq 15$ or $R \leq -15$ we will reject the hypothesis of independence at the 1% level of significance.

Since $R = 14$ in our example and if we choose the 5% level of significance we will reject the hypothesis and conclude that the X's and Y's are associated in some manner. If we choose the 1% level of significance, we will accept the hypothesis of independence because we failed to reject it.

This test practically ignores the data at the center of the scatter diagram and focuses attention to the points at the edges of the scatter diagram. The test does not measure the degree of association as does the correlation coefficient.

For this test one assumes that the populations are continuous so no ties should occur. If ties do occur they will require special attention. If two (or four) points should lie on Y median line, then move the first point up or down, as determined by the toss of a coin, and move the other point in the opposite direction. The same thing applies for points lying on the X median line. If in constructing one of the L lines one meets with a set of tied observations, then draw the line through these tied points but take the contribution to r to be

$$\frac{n_s}{n_o + 1}$$

where n_s is the number of points in the tied set which are on the same side of the median line as the initial point, and n_o the number of them on the opposite side.

Comments and conclusions

For data that are measured in the nominal or ordinal scale one usually uses nonparametric tests.

For small samples nonparametric methods are usually easier to compute than the parametric methods.

For large samples it may be easier to use parametric tests since it can be tedious to rank data unless one has sorting equipment.

For data that are expensive to obtain one would want to use the most efficient test available.

For nonparametric tests one does not specify the functional form of the frequency function.

The three tests illustrated in this paper are only a few of the tests available. It is hoped that the people working in statistics and Quality Control will become better acquainted with these nonparametric techniques as they could be most useful to them.

References

1. Siegel, S., 1956 Nonparametric Statistics New York: McGraw-Hill

2. Dixon, W. J. and Massey, F. J., 1957 Introduction to Statistical Analysis New York: McGraw-Hill

3. Walker, H. M. and Lev, J., 1953 Statistical Inference New York: Henry Holt and Company

4. Mood, A. M., 1950 Introduction to the Theory of Statistics New York: McGraw-Hill

5. Wallis, W. A. and Roberts, H. V., 1956 Statistics A New Approach Glencoe, Illinois: The Free Press

INITIATING QUALITY CONTROL AND QUALITY STANDARDS FOR NEW PRODUCTS

William P. Cloyes

Titanium Metals Corporation of America

Introduction

The problem of establishing quality standards in a new industry is a difficult one, as many here will agree. In the first place, basic information is lacking which is needed for accurate application of modern quality control methods. Particularly, there is no accumulated knowledge from past experience in the following two fields:

1. Process capabilities
2. Customer requirements

Lack of information concerning process capabilities and customer requirements makes application of modern quality control methods a matter of continually aiming ones testing and inspection methods towards establishing proper quality standards. The aircraft industry as customer has been very helpful and cooperative in setting up attainable specifications and evaluating the quality requirements of this new metal, titanium.

The purpose of this paper is to discuss in detail the quality control developments within Titanium Metals Corporation of America during the past two years in establishing control and standards together with the accomplishments made under this program.

Background of the Quality Control Program

As many here know, titanium is a new metal to American industry and America far outstrips the rest of the world in production of this metal. Titanium metal is presently primarily used in the manufacture of jet aircraft--both for engine and airframe parts. It is particularly valuable in these applications due to its high strength at elevated temperatures up to 1000 degrees F, its extraordinary corrosion resistance, and its low weight compared to steel (specific gravity 4.54 vs. 7.92 or approximately 43% lighter).

Titanium Metals Corporation of America was formed in 1950 as a joint venture by the National Lead Company and the Allegheny Ludlum Steel Corporation. TMCA has been fortunate in this background, as National Lead brought many years of experience in dealing with titanium chemicals and Allegheny Ludlum many years of experience in rolling and processing high-strength alloys. Titanium Metals Corporation of America was the first fully-integrated organization in the industry, with operations ranging from titanium ore through sponge production and ingot making, to the production of all mill products - billet, bar, sheet, strip, plate, wire, extrusions and tubing. The fantastic growth of the industry can be seen from the production figures of 3 tons of sponge in 1948 against approximately 14,500 tons of sponge and approximately 5,300 tons of finished mill products in 1956.

The period since 1950 has especially been a period of rapid expansion, culminating the past two years in shipments quadrupling previous production records. This production increase has placed a continual heavy load on our quality control program, not only to maintain a uniform quality level but also through additional quality controls and research, to improve our product to satisfy the customer's increasing quality demands.

Establishment of Quality Control Program

TMCA's first step in developing its quality control program was to make a survey of all of its operations from the raw materials through to the finished products. The report from this quality control survey provided a plan for long-range development of modern quality control methods in our company. The report suggested steps toward developing quality standards and quality control procedures throughout our operations and our program has basically followed these recommendations.

TMCA's program has been continuously under the direction of a Quality Control consultant. At the first of the year we created the new position of Quality Control Manager in our organization. Mr. David J. Ausmus, former Chairman of the Youngstown Section ASQC, has been selected for this position.

Process Capabilities

As stated above, our initial quality standards had to be set without accumulated knowledge of process capabilities, which would have permitted the setting of scientific specifications for processes and products. Orders had to be taken for new products, processes had to be developed, and products had to be produced--all without detailed knowledge of process capabilities.

During the past two years, however, we have made careful studies of all of our operations and have developed substantial information concerning quality standards for our processes and products. These are now serving us well, and are explained below.

Customer Requirements

As new products have been developed, no information was immediately available concerning the needed quality standards of customers over and above the stated specifications. We have found several properties chemical, physical, and visual, of titanium metal which were not covered in the original specifications but which had an important effect on the customer's use of the metal.

Even though these characteristics were not covered by specifications, the customer needed protection and we have developed quality standards and quality controls to cover these characteristics.

Quality Standards in Ingot Manufacture

The most important quality control problem in ingot manufacture is that of making tests and inspections which will correlate some considerable time later with specification quality of the finished

products. For example, customer specifications of finished products are usually in terms of tensile strength, yield strength, elongation, and reduction of area. However, titanium metal in ingot form cannot be evaluated accurately in these terms. Therefore, the relationship of associated ingot tests, such as Brinell hardness, needed to be closely established.

We have made correlation studies of such relationships and have established quality standards for ingot Brinell hardness that permit us to know in advance how the ingot will work out in finished products. The control of ingot Brinell hardness measurements is maintained by standard X & R charts.

Information from these control charts is summarized in monthly management reports and distributed to interested persons throughout our operations.

Another important problem in ingot quality control is maintaining physical uniformity or homogeneity of the ingot, such as avoidance of pipe in the ends of the ingot. Pipe is a void situation of some length in the center of the ingot ends. We have been able to set up tests for pipe in individual ingots at Henderson, Nevada and have correlated these with pipe losses in finished products at the mills in the East, through use of statistical methods. Such correlations have proved of tremendous value and have enabled us to develop continuous control procedures at Henderson which have substantially reduced our pipe losses in finished products. Frequency distributions of X & R are also compiled for alloying and trace elements found in the metal.

Quality Standards In Mill Operations

We have developed sampling controls for determining process capabilities of gauge in flat-rolled products and have maintained these on standard control charts. These have proved valuable in considering what tolerance limits may be guaranteed on close gauge orders. Standard X & R charts are also maintained for control of physical properties and a summary tally is issued to management monthly.

Similarly, we have used the methods discussed in "Quality Control Techniques for Establishing Industrial Standards" (1) in developing quality standards for visual characteristics of our products. These are still in process of development.

A testing program to evaluate the testing results of TMCA and Allegheny Ludlum laboratories is also in progress and this data will be evaluated statistically.

Current Developments

We have found that the methods of statistical quality control have improved communication of quality information among all interested parties throughout our organization. This has enabled us to take more prompt action when quality standards are violated anywhere in the processes.

From the above discussion, it will be noted that considerable progress has been made in our efforts to develop close quality standards for our products and a quality control system that will continually meet such standards. We have a long-way to go toward our ultimate goal, but feel that we are on a solid road of progress toward that goal.

Reference (1) Wareham, Ralph E., "Quality Control Techniques for Establishing Industrial Standards", National Convention Transactions 1955, American Society for Quality Control, 359 - 364.

DESIGNED EXPERIMENTS IN INDUSTRY

By Dr. H. C. Hamaker
Philips Research Laboratories, Eindhoven, Netherlands
and
Rutgers University

A number of lantern slides will be shown illustrating case histories of the successful application of designed experiments in industry. These examples then form the basis for a more detailed discussion of some basic principles underlying the use of experimental designs in an industrial environment. It is in particular pointed out that textbook treatments of the subject are far from satisfactory and that important changes are needed before experimental designs will bear full fruit in industry.

Usually a large number of factors have to be considered and the true art of experimentation exists in deciding which of these factors should be varied in the first place and which be kept constant or studied in later experiments. The experimenter then proceeds by a series of successive experiments which jointly provide the answer to this problem. This aspect is neglected in statistical textbooks where the analysis of a successfully designed experiment is discussed, not the true problems involved in the design of the experiment itself.

In the literature there is far too much emphasis on analysis of variance and the testing of hypotheses. In many cases properly arranged tables of averages reveal the result of an experiment with sufficient clarity to serve as a basis for further experimentation. More sophisticated statistical analyses are uncalled for; they may be used by the statistician to check the validity of the main conclusions but they should not find a place in his report to the experimenter, or at best should be placed in an appendix.

Technical and statistical significance are not sufficiently distinguished. In a complex analysis of variance very many effects including first and second order interactions may turn out to be statistically significant, but technically they may be too small to be worth considering. In such situations we should pick out the large effects and report on these, pooling all the smaller effects into a single pooled residual. Too much detail will confuse the technologist and before something has been done on the large effects the small effects are not worth worrying about.

It is often emphasized in textbooks on statistics that we can only test our preconceived models set up before the experiment was made. The experimenter can never accept this point of view. To him the numerical determination of the parameters of a preconceived model may be of some value indeed. But it will often be of much greater interest if the experimental results suggest that the preconceived model was wrong and that an all-together different model has to replace it. The most successful work of the statistician in industry is where he can show that such a change of model is needed. This will invariably lead to much more effective and efficient methods of experimentation.

Statistical textbooks always show examples of designed experiments where the analysis previously explained does apply. In practice, however, many designed experiments yield data from which it is at once obvious that the standard methods do not obtain and that some different method of treatment has to be adopted. The success of a statistician in industry will in a large measure depend on his ability to recognize these situations and deal with them effectively. They should be included in the literature.

The full impact of the "Design of Experiments" will only be realized when the subject is taught in the universities as part and parcel of experimental science and not as a branch of statistics, itself a branch of mathematics. At present progress is hampered by the biassed outlook of the statisticians who see the subject too much from the standpoint of probability theory.

THE TEAM APPROACH IN QUALITY CONTROL INVESTIGATIONS

Charles A. Bicking
The Carborundum Company

Introduction

Eighty-eight per cent of the articles published in "Industrial Quality Control" have one author, ten per cent have dual authorship and only an unimpressive two per cent have more than two authors. (a) This might be taken as an indication that Quality Control is a lone-wolf operation. There may well be some basis for this inference because in the early days, in particular, a torch was being carried by lone individuals in many an organization and the very existence of a quality control activity was due to their missionary zeal. Even now, formally organized quality control departments employing more than a very few engineers, except in large organizations, are not the rule. Although the educational activities sponsored and encouraged by the American Society for Quality Control have reached tens of thousands of people in industry, many are constrained to apply the techniques as a part-time assignment and without formal organizational recognition of their quality control function. In spite of the dependence on individual initiative, however, the team approach has been used for a long time.

The margin between failure or success may in many instances have been the degree to which teamwork among the quality control engineer, technical people and production personnel was accomplished. Relatively little has been published about the use of the team approach, so that one might conclude that it was a natural rather than a conscious use on the part of many of those who succeeded.

In a somewhat newer management technique, Operations Research, the concept of the team is one of the essential characteristics. There are so many similarities between the conduct of Operations Research and Quality Control that it seems desirable that more light be shed on the part teamwork has to play in the latter technique.

The Essential Characteristics of a Team

A successful team is remarkable on several counts. It brings together a number of people, each of whom has something different to contribute to the solution of a problem. As individuals, the members of the team recognize that each

---

(a) Immediate source of information: Mr. Ervin F. Taylor in a tentative proposal of the Bibliography Committee, ASQC, for a Bibliographical Service, February 1957.

has something special to contribute to the success of a job. Mutual respect and a recognition of the talent and special attributes of the other members provide a basis for effective cooperation. The work of the team is so organized that the various members look at the same phases of each problem from a different point of view.

In order to be recognizable as a team, the group working together must have a common goal. The members must have a willingness to cooperate. It helps if they have had earlier mutual experiences, but since there has to be a beginning sometime, even a group of strangers can be welded into a team under the right conditions. A cardinal condition is that each one recognize that no one is competing in another man's territory.

Good members of a team must be willing to try new combinations of ideas and things and to accept new approaches or attitudes. There must be a willingness to accept the initiative in following up on plans agreed upon by the team. Each man must be able to work under self-direction after a general outline has been agreed upon.

Cooperation must be just as strong at planning stages as at stages of execution or evaluation of a plan. Members of the team must be able to take and give constructive criticism impartially. Each must have the ability to accept compromise, to agree on definitions, on the setting of schedules and on the necessary limitation of individual activity. A mark of a good team worker is a willingness to share fairly the credit for a task well done.

These characteristics are just as important whether the teamwork has come about naturally and artlessly or whether it has been the result of a coordinated plan. With the advent of professional teams in the consulting field as well as on the staffs of large manufacturing organizations, some additional general characteristics have become recognizable.

Such teams have the support of top management and are given considerable freedom of action. The level of maturity and professional ability is very high. Work of great responsibility is attractive to capable people. The members of these teams are hand-picked. Because of the respect that they command, the team often has much influence and can evoke the authority to act when necessary. The very air of authority and assurance makes accomplishment of difficult tasks possible.

There is great moral as well as technical force behind a team that has a systemmatic program. There is a great advantage in having a system for assignment of responsibility for action and in maintaining a strict check-off as action is taken. Experience has led to the breaking down of complex projects into parts of manageable size and to the assigning of each part to a specially qualified team, often of not more than three to six men.

It may be instructive to recall some examples of application of the foregoing principles from past experiences which in retrospect appear to have been characterized by good teamwork, to describe the exceptionally productive teamwork between Army Ordnance and the Bell Telephone Laboratories during and since the Korean War and to present some current developments in the team approach in process quality control.

Artless Uses of the Team Approach

Teamwork on a Pilot Plant Problem

A number of years ago, the general management of a chemical company was providing a new and specialized advisory service to the plants in the form of assistance in the introduction of statistical methods of quality control. The quality control staff had recognized, however, that in the chemical industry the study of analytical precision and the statistical design of experimental work should accompany or precede attempts toward process control. The company quality control engineer was therefore spending some time at a plant laboratory in order to induce chemists and development engineers to expand their use of statistical methods.

Also at the plant at the time was a staff process engineer who was working with one of the project development chemists in the laboratory on a pilot plant operation. At the suggestion of the project engineer, the process engineer and the quality control engineer met with him at the laboratory one morning to discuss the possibility of a chart method for control of the pilot plant. (b) The three men had known each other for some time. The two from the home office staff, although they did not ordinarily work as a team, both had considerable experience in helping solve production problems and both had expressed willingness to work together with interested project engineers on current development problems. This team met only once for a very brief time in connection with problem mentioned and thereafter carried out a plan and achieved results very rapidly. The details of the plan have been published (1) and need not be repeated here. However, a brief description of the method of operation is essential.

Data already at hand were used almost exclusively. A scatter diagram for the per cent conversion against one of the control characteristics was made using results of a special run. The resulting scatter confirmed what was known about the reaction. A curve was fitted to the plotted points and plus and minus limits were drawn about the curve so that 95 per cent of single determinations should lie within them if proper processing conditions were being maintained.

(b) Members of the team were respectively Mr. John Drew, now with the Crosby Chemical Company, Mr. Homer A. Smith, now with the Minerals and Chemicals Corp. of America, and the author. All three were formerly with the Hercules Powder Company.

Operators were instructed to raise or lower temperatures depending on position of each routine test with respect to the control limits. They found the method was much easier and quicker than the former tedious method of control. The project engineer began to experiment with different temperatures, pumping rates, frequencies of catalyst change, and numbers and arrangements of reactor vessels. It was soon discovered that best conditions were produced by using the reactor with fresh catalyst ahead of the reactor with partially spent catalyst, in contradiction of accepted procedures. It was estimated that the results were achieved six months earlier than would otherwise have been possible and that the total cost saving was ten per cent.

One might speculate on what elements in this situation caused it to be remembered as so remarkable. It may be simply that all the essential conditions for successful team work were fully satisfied.

Teamwork on a Plant Process Problem

Another example of teamwork dating back to the early days of statistical quality control in chemical processes has been described recently. (2) This has to do with the introduction of control charts for the first time in a plant where supervision had almost no prior experience with statistical methods. The staff quality control engineer traveled to the plant. He met with the plant superintendent, the control chemist, and a development chemist to discuss what should be done. (c) During a two day visit, this team reviewed preliminary charts showing prior data plotted between specification limits, studied the process to select appropriate control points and outlined a control chart program for three processes and a precision determination program for the principal chemical analyses involved.

During succeeding months, the members of the team permanently stationed at the plant worked steadily on achieving process control with occasional advice from the quality control engineer. Within two months, the control laboratory had made replicate tests on homogeneous samples to arrive at precision figures for all the principal control tests. Control charts began to show a steady improvement in uniformity. For one process, the control chart was of assistance in locating an error in calibration of the feed tank for a raw material. Correction of this defect led to greater ease in handling the process and greater uniformity in product.

At the end of the first six months the uniformity of the various characteristics of one product had been improved anywhere from 43 per cent to 54 per cent with an accompanying

---

(c) This team consisted of Mr. Earl Radant, Mr. Jesse Loucks, Dr. John W. Smith, and the author, respectively. All were employees of the Hercules Powder Company.

reduction of 6-1/2 per cent in processing time. For another product, uniformity of one characteristic improved 60 per cent and other characteristics were being brought into control. A third product was so much improved that it was possible to produce a USP grade within specifications 100% of the time.

Credit was consistently given to the fine teamwork by the group put together at the plant to handle this project and their fine cooperation with the quality control engineer.

Teamwork on an Industry-wide Basis

On occasion, our industrial system gives rise to opportunities for team collaboration of an unusual sort. An example is an informal team that had a considerable impact on the acceptance of statistical quality control in the pulp and paper industry. (d) These three men, a research chemist, a technical service engineer, and a quality control engineer, collaborated over a period of years. All were employees of the same company but were in different divisions. After assignment to a common project brought them together, they continued to collaborate on a spontaneous basis.

This collaboration consisted of educational activities within the plants and laboratories of the company and in the plants of customers. It involved the planning, conduct, and analysis of mill trials of company products. It involved introduction of quality control techniques to the plants of the company and of experimental techniques to its laboratories. It involved work in the Statistical Committee of the Technical Association of the Pulp and Paper Industry. It involved the preparation of bibliographies and the distribution of technical material on quality control. It involved personal selling of quality control principles on many private and social occasions. It involved mutual review and criticism of plans, analyses, and reports of testing. It involved publication of significant contributions of the literature. It involved collaboration in teaching short courses under the auspices of the Technical Association.

The feeling of identity of interests, a high mutual regard for each other's various talents, and an unselfish, almost disinterested sharing of credit led to results none of the three could have achieved alone.

Teamwork Perfected for National Defense

Quality Control of 105 mm. Ammunition

The tremendous job done for Army Ordnance during the Korean War by the Quality Assurance Department of the Bell

---

(d) This team consisted of Mr. Richard T. Trelfa, a Fellow of ASQC, now Plant Superintendent, Watervliet Paper Co., Mr. Harris Ware, now Manager of Technical Service, Beveridge Paper Co., and the author. All were then employees of the Hercules Powder Company.

Telephone Laboratories is well known. (3) The project for Quality Control of 105 mm. Complete Rounds resulted in successful production of an initial large lot of over 149,000 rounds within which the dispersions of muzzle velocity and range were fully as low as for the very much smaller lots, of the order of 30,000 rounds, (e) produced prior to the beginning of the project.

What has not been so well publicized is the fact that the success was due to the remarkable teamwork that existed between Bell Laboratories and Ordnance personnel. In the words of Mr. George D. Edwards, Project Engineer for Bell Laboratories,

> "During a considerable part of World War I and throughout practically the whole of World War II, I worked with various branches of the armed forces in a capacity somewhat similar to the one in which I now find myself. Never before have I experienced the sort of cooperation, and the genuine desire and ability to cut red tape, and to do everything possible to forward the job, that the Ordnance Ammunition Center has shown."

This teamwork was the result of a new concept in organization. The contract called for Bell Laboratories to provide a skeleton organization of quality control engineers and for Ordnance to parallel this with an organization of 'opposite numbers' in each case. There were two reasons put forward for this provision: first, to keep the cost of the contract as low as possible; and second, to train Ordnance personnel as rapidly as possible, so that they could take over full responsibility for the work at the earliest possible moment. Actually, the 'opposite number' concept evolved as an extremely valuable organizational principle, and has been applied successfully to later Bell Laboratories-Army Ordnance projects and is being used successfully in industry.

Throughout the 105 mm. project the team was favored with constant help and cooperation of the regular Ordnance supervisors and personnel as consultants, advisers, and as instructors in ammunition to Bell Laboratories engineers. Also, Ordnance contractor personnel from American Safety Razor, Chevrolet Shell, General American Transportation, Goodyear Engineering, Kelsey-Hayes Wheel, Malleable Iron Range, and Thor Corporation, were most helpful at all times.

Quoting Mr. George Edwards again,

> "Process quality control is frequently an elusive sort of thing. We (the Bell Laboratories members of the team) have certain standard tools,

---

(e) Lot sizes were commonly much smaller than this during World War II, being then of the order of 3000 - 4000 rounds. Ordinary quality control techniques had increas- the lot size to about 30,000 before Korea.

such as control charts, analysis of variance, etc., which we can apply. But if none of them work, we still have to get the answer, and we are forced to fall back on 'ingenious devices' for doing so. That is when Ordnance knowledge and experience are able to make a very special contribution to the job."

Quality Assurance Program for the Nike Missile

This use of opposite numbers worked out so very well that when, in 1954, Bell Laboratories undertook the assignment of developing quality assurance procedures for the Nike I guided missile, it was again included as a contractual requirement. (4) Although the far-flung organization required to cope with a production effort of the size of this one made difficult the keeping of an effective team in the field, a major effort was made throughout the contract to maintain the balance of 'opposite numbers' on the team. This effort was fully vindicated.

Another excellent example of the use of the team approach, although not new, (f) was the practice of using the committee technique in making quality surveys of various components and sub-assemblies of the missile. The Survey Committee usually consisted of three members, authorized to represent respectively the quality assurance agency, the inspection agency, and the supplier. For example the first quality survey on the Nike Project involved the ballistic test of the booster. The committee consisted of a Bell Laboratories quality control engineer, and an Ordnance 'opposite number' from the project office (for the Quality Assurance Agency), an engineer from the plant manufacturing the booster (for the supplier) and the Ordnance resident inspector from the plant (for the inspection agency). (g)

The operation of a quality survey committee has been described by Mr. E.G.D. Paterson, who supervised the Nike Quality Assurance project. (5)

Each member of a survey committee was responsible for those factors affecting quality which were within his purview. For convenience, the subject areas were divided as follows:

A. Quality Assurance Agency

1. Contracts, specifications and drawings
2. Examination of a sample of current product
3. Inspection procedures
4. Inspection results and surveillance results
5. Service complaints
6. Review of quality standards

(f) A quality survey procedure employing a committee or team approach had been in successful operation in the Bell Telephone System for something over 25 years.

(g) This team consisted of Messrs. G.R. Gause, N.C. Gardner, A.F. Giacco, and M.C. Willard, respectively.

B. Inspection Agency

1. Inspection procedures
2. Gage and test set calibration and records
3. Service complaints

C. Supplier

1. Manufacturing drawings, handbooks, and layouts
2. Manufacturing facilities, including gages and test sets
3. Manufacturing inspection procedures
4. Manufacturing difficulties
5. Raw materials inspection

The committee survey technique as used by the Bell Laboratories has an impressive record of success.

Artful Use of the Team Approach

In a recent paper, (6) General Leslie E. Simon suggested that, particularly if operations are large or complex, launching of a statistical quality control program "can be done most economically by assembling a crew of experts to design, install and initially operate the S.Q.C. system." He suggested, also, that "concurrently, 'opposite numbers' drawn from existing personnel or obtained by recruitment, must be assigned to training so that they can take over the operation from the experts."

Continuing to quote General Simon,

> "The installers of a quality control system must secure teamwork between themselves and operating personnel to devise alterations in process and in process inspection for the purpose of bringing each stage of the process to a state where it will respond to statistical control. It is very important that the installers of S.Q.C. and the permanent personnel be friends.
>
> "The exercise of ingenuity both in statistics and in engineering is necessary to bring a process under statistical control. Changes must be made in the flow of product so that each increment arising from a different source can be checked for quality before it is allowed to mix and lose its identity. New inspection points must be set up for each check point. The most knowledgeable people associated with the process must be brought together to seek out causes for lack of control and devise ways and means for their elimination."

This is the general concept of staff quality control organization now being carried out at The Carborundum Company. In the early stages, the size of the team has been variable both because new members of the quality control group have

been in training and because some missionary work has to be done before operating divisions see the problem clearly enough to make permanent assignments to the team. Eventually, it is expected that each study involving a major process will call for a team of perhaps three members of the quality control staff; a senior quality control engineer plus two junior engineers or quality control technicians. It is expected that operations will appoint two or three engineers or technicians to each team.

The senior engineer is captain of the team and is responsible for all phases of planning and execution of the program for the assigned process. Assisted by members of his team he makes a verbal report of progress at a weekly staff meeting. He organizes a brief, factual progress report at the end of each month. This is combined with other team reports in a report by the Quality Control Manager to the Vice President in charge of Research and Development. The team captain, or the responsible members of his team, write technical reports including conclusions and recommendation at the completion of each major phase of the program. These reports are checked with the operating division members of the team and with operating supervision and are issued jointly. Appropriate actions are assigned either to the Quality Control Branch or to the Operating Department. At the completion of the study of a process, a final report or reports are written and the process control system established is turned over to the cooperating division for routine operation. The concept calls for routine surveillance of each completed program to assure its continued effectiveness.

Conclusion

Several types of team approach to a statistical quality control program have been mentioned. First, there is the frequently found approach which relies upon individual initiative. A one-man quality control department must use this approach or vegetate in an ivory tower. Secondly, there is the use of 'opposite numbers' in a team made up more or less equally from a consulting, or staff group, and an operating group. This has been very useful in stepping up effort on particular programs and carrying them to most rapid conclusions. Finally, there is the conscious effort to exploit the team approach in the operation of a staff quality control function in a multi-division manufacturing organization. This promises to maximize the impact of quality control in breadth and depth.

There are some difficulties involved in the successful use of teams that are formally organized and strictly assigned to a project, particularly if the project is of limited duration. It is difficult to put together a team if the future of the assignment is not completely clear. Individuals may be doubtful about what happens upon completion of the project. On the Bell Laboratories-Ordnance projects, difficulty was experienced in inducing people to move to a location geographically distant from home base on a temporary and even uncertain

basis. This was more difficult when Ordnance people moved to the Laboratories to take part in the Nike project than when the Laboratories moved some of its staff to the Ordnance Ammunition Center for the 105 mm. project. Both of these projects had the additional disadvantage that the production operations concerned were scattered throughout the country.

There is the difficulty of finding suitable people in operating divisions which have never used S.Q.C. to assign to the teams. Usually, available people also have routine duties which either must be neglected or at least temporarily reassigned.

There is the effort attendant to training someone from the operating group to take over the procedures when the team completes its work and pulls out.

Finally, there is the increased problem of communications when larger numbers of people are directly involved in a project.

These disadvantages are more than off-set by the many advantages of the team approach. Once communication is established, results are achieved much faster than would be possible otherwise.

Industrial relations are improved by the feeling of belonging to the team. The use of the term 'opposite numbers' is perhaps unfortunate because exactly the reverse of opposition is implied by teamwork. Perhaps an improvement could be made by referring to the team as made up of "complementary members" from the quality control staff and the operating division.

Another point in favor of the team is that experience has demonstrated that it is the most effective way of operating quality control as a staff service.

Perhaps the greatest advantage of all lies in the fact that once a procedure has been agreed upon and put in operation it has a greater likelihood of sticking since those most closely concerned, the operating people themselves, have had a part in the whole endeavor.

References

1. "Operational Research in a Chemical Pilot Plant", Charles A. Bicking. Handbook of Industrial Engineering and Management, Ireson and Grant, Editors, Prentice-Hall, Inc., 1955. pp. 1028-1031.

2. "Quality Control in the Chemical Industry", Charles A. Bicking. Proceedings Fourth Annual Forum, Toronto Section, ASQC, Toronto, March 9, 1957.

3. Minutes of Progress Review Conferences, Quality Control 105 mm. Complete Round, Bell Telephone Laboratories, Inc. 1951-52.

4. Minutes of Progress Review Conferences, Quality Assurance Nike I Round (Unclassified), Bell Telephone Laboratories, Inc., 1954-56.

5. "An Over-all Quality Assurance Plan", E.G.D. Paterson, Industrial Quality Control, Vol. 12, No. 11, May 1956, pp. 32-7.

6. "On Initiating a Quality Control Group in an Existing Organization", Leslie E. Simon, Western Regional Conference, ASQC, 1956.

STATISTICAL APPLICATION IN THE CONTROL OF HIGH SPEED FILLING MACHINES

Francis X. Donohue
The Nestlé Company, Inc.

The control of weights in the modern, high-speed filling line is tremendously important, particularly in consideration of (a) economic factors, (b) sound commercial and consumer relationships, and (c) continuous conformance to legislative requirements pertaining to package weights.

I do not believe that there is any need to discuss these factors in great detail as they have been thoroughly and adequately covered elsewhere. It is sufficient to say that economic factors are usually the most compelling motives for close control of filling operations simply because the annual losses due to excessive overfill or "give away" can easily represent many thousands of dollars. Filling losses are an easy target for modern cost analysis techniques which seem to elevate such excesses to bold relief. The compelling need for weight control of package filling lines presents a real challenge to the quality control engineer.

In filling control, the basic goal might be defined as that of establishing the average filled weight at the lowest level consistent with the declared label weight and good commercial practice. In our experience, package weights usually conform to normal distributions. Consider the very simple sketch shown in figure 1 in which the distribution of package weights is graphically shown in relationship to the declared label weights.

Federal regulations covering package weights require that the average weight of a lot be at least equal to label weight and that there be no unreasonable shortage in any one package. They do not require that each individual package be of label weight. The principal exception to the above is found in those instances where the label declaration specifies a minimum weight in which case all packages must contain the declared weight. This latter situation occurs relatively infrequently, however.

It is apparent that one way of establishing a target average for a given filling operation would be in terms of (a) a tolerance specification for the percentage of underfilled individuals issuing from the line that are acceptable; and/or (b) the maximum shortage in any one individual that is acceptable. Either or both of these requirements may be zero. Having established these tolerances, one end of the distribution becomes fixed by the limiting factor. The average overfill then will be determined principally by the filling variability. Having established one end of the distribution, the only practical measure for decreasing the average overfill, without changing the specification, is the reduction of filling variability. It is for this reason that a thorough investigation and evaluation of filling line variability is recommended as the first step in approaching a control problem of this sort.

Digressing for a moment, I wish to make it perfectly clear that the definition of acceptable weight tolerances for a particular package

Distribution of Net Weights
Relationship of Decreased Filling Variability to Overfill

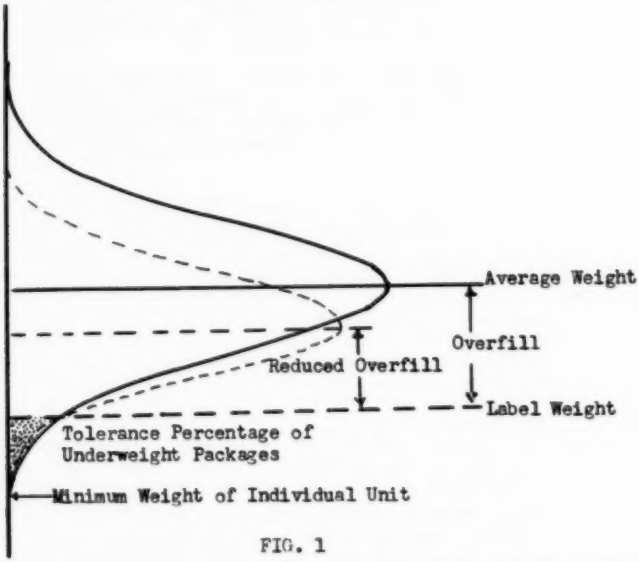


FIG. 1

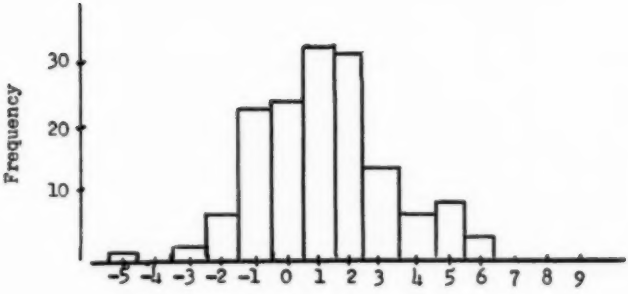
is a subject for management decision. The variability of the package weights depends entirely upon the capabilities of the filling equipment used and the characteristics of the material being filled. Therefore, specific definitions of acceptable net weight variability must be resolved in the light of the circumstances which comprise a particular situation.

The control problem naturally resolves itself into two parts. First, one must estimate the process variability and, secondly, a technique must be devised for the control of the line that is sufficiently rapid and sensitive to match the capacity of the filling machine under consideration.

Very often, variability patterns for filling operations are complex and confusing. This is particularly true of multi-head machines where overall estimates of a filling variability are in reality a combination of variances that include the following factors.

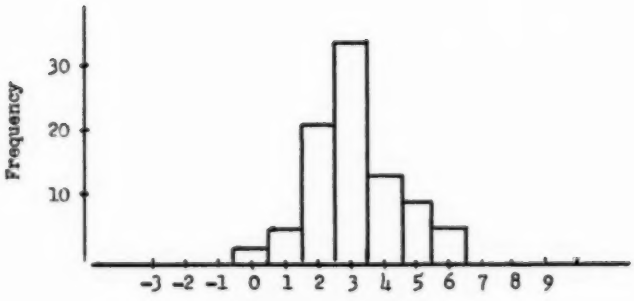
(1) Product variability
(2) Variance due to a difference in the average weights filled by individual heads
(3) Lack of homogeneity of variability in distribution of individuals from the several heads
(4) Sampling error

One of the most useful techniques in the evaluation of filling line variability is the replicated factorial analysis of variance. In this connection, 100% sampling for very short periods of time has been most useful. To illustrate, a typical data analysis is shown in figures 2, 3, 4 and 5 attached. All of these figures refer to large scale filling



Frequency Distribution
of Net Weights Filled on Unit A

FIG. 2



Frequency Distribution
of Net Weights Filled on Unit B

FIG. 3

FILLING CONTROL TEST DATA*

VARIANCE ANALYSIS PRODUCTION FROM UNIT A

Head	Test 1 Sample 1	Test 1 Sample 2	Test 2 Sample 1	Test 2 Sample 2	Test 3 Sample 1	Test 3 Sample 2	Average
1	-1	-1	0	0	-2	-1	-0.8
2	2	0	2	0	1	1	1.0
3	0	2	2	1	3	-1	1.2
4	0	1	2	0	0	2	0.8
5	0	1	3	0	4	3	1.8
6	0	-1	3	1	1	2	1.0
7	-1	-2	1	0	2	-2	-0.3
8	5	4	5	5	5	4	4.7
9	4	5	4	4	6	3	4.3
10	-5	-3	-1	-1	-1	-1	-2.0
11	0	-1	-1	-1	1	0	-0.3
12	2	2	0	-1	3	2	1.3
13	3	3	4	2	6	6	4.0
14	2	1	2	1	2	1	1.5
15	1	1	-1	1	1	1	0.7
16	2	2	1	2	2	2	1.8
17	0	0	0	0	2	1	0.5
18	3	5	5	5	3	3	4.0
19	1	0	3	1	-1	-2	0.3
20	-3	-2	-1	-1	-1	-2	-1.7
21	-1	-1	1	1	1	1	0.3
22	-2	-1	-1	0	0	0	-0.6
23	3	1	2	1	2	2	1.8
24	2	3	2	2	1	0	1.7
Av.	.75		1.25		1.38		1.13

FIG. 4

*All data coded

FILLING CONTROL TEST DATA*

VARIANCE ANALYSIS PRODUCTION FROM UNIT B

Head	Test 1 Sample 1	2	Test 2 Sample 1	2	Test 3 Sample 1	2	Average
1	2	2	2	2	2	-1	1.5
2	2	2	1	3	4	4	2.7
3	1	1	2	2	1	1	1.3
4	2	2	2	1	2	2	1.8
5	-1	1	5	4	4	2	2.5
6	5	5	4	4	5	4	4.5
7	2	2	2	2	2	2	2.0
8	3	2	1	5	2	3	2.7
9	2	2	1	1	3	3	2.0
10	1	2	3	2	2	3	2.2
11	4	3	4	3	3	2	3.2
12	1	1	3	1	0	2	1.3
13	0	0	3	2	2	0	1.2
14	1	1	1	0	1	2	1.0
15	1	2	3	1	1	1	1.5
Av.	1.80		2.33		2.10		2.08

FIG. 5

*All data coded

operations and you will note that the data is coded and therefore has no absolute significance. I should also mention that figures 2 and 4 represent typical production from a high-speed filling line which operates at speeds in excess of 200 units per minute (Unit A). Figures 3 and 5 describe the performance of a similar machine (Unit B) that operates at speeds of approximately 150 units per minute.

The data used in the construction of these figures are the net weights of three series of samples drawn from each machine. Each series consists of two consecutively filled samples drawn from each filling head. The samples were drawn at randomly selected times over a three-hour period. During the test periods, the machines were controlled by an operator who drew frequent individual random samples and made such adjustments as he felt were necessary. No control charts were used by the operator and his sole criteria of control was the requirement that the average weight should be maintained at the lowest level consistent with producing all units of label weight.

Referring first to the frequency distributions shown in figures 3 and 4, it is apparent that the variability in production from Unit A is substantially larger than is found in the production from Unit B. The calculated standard deviations are, in fact, 2.0 and 1.3 respectively. These estimates of filling variability could be used as a basis for construction of control chart operating limits. However, as previously noted, overall estimates like these are actually a combination of several variances, some of which may be controllable. To illustrate, the data from which figures 2 and 3 were constructed are also shown in the well known factorial variance analysis format. Figures 4 and 5 are actually two-way tables of the data and reflect in themselves something of the true nature of the overall variability. The detailed calculations required for the variance analysis are not shown. However, they are relatively simple and can be found in most texts on the subject. A comparative summary of the mean squares is shown in figure 6.

SUMMARY OF VARIANCE

	Unit A		Unit B	
	Mean Square	Dif.	Mean Square	Dif.
Between Heads	18.83	23	5.14	14
Between Tests	5.25	2	2.10	2
H x C	1.88	46	1.67	28
Residual	.79	72	.77	45
Total		143		89

FIG. 6

Perhaps the most interesting result of this test is found in the residuals which are substantially the same. In the case of Unit A, there is a relatively large interaction, which subsequent tests revealed to be somewhat abnormal. This was probably due to the fact that several of the filling heads became partially plugged during the test. With regard to the between test main effect, Unit A evidenced significance at the 5% confidence level. This is simply interpreted to mean that as a result of variations in the physical characteristics of the powder during the fill period, the overall machine average varied significantly with respect to the inherent variability of the machine to produce uniform weights. It is also an indication of the inadequacy of the control systems employed at the time of the test. The between head main effect in both units was highly significant. It is quite obvious that the principal problem in controlling these units is the reduction and variability of the average weights issuing from the individual filling heads. Motivated by the results of tests such as these, corrective measures were undertaken which have substantially reduced this variability, particularly on Unit A.

Having established a reasonably clear picture of filling variability, two conclusions can be drawn. It is apparent that the two machines are characterized by approximately the same inherent capacity to fill uniform weights assuming, of course, that the fill from the individual heads can be uniformly maintained at essentially the same level. Secondly, the residuals in themselves provide a preliminary estimate of the theoretical capacity of the machines to fill uniform weights. These estimates could be used as a basis for calculating operating control limits.

In developing the specific control technique, consideration must be given to other factors. The product is packed in glass which makes gross weighing procedures cumbersome and difficult but not impossible. This means then that net weighing has to be resorted to. The control of machine average can be accomplished in several ways. A control based on individual samples is rather impractical because the time required for evaluating enough samples to get a reasonable estimate of machine average is excessive when the machine average is shifting. The target average fill can be set very close to the label weight without undue risk of producing underweight units providing the control technique is sufficiently sensitive to detect incipient shifts in average fill which are occasionally encountered.

Another factor that has an important effect on the control of filling is concerned with the accuracy, sensitivity and speed of the balances used. Actually, this is a subject in itself which cannot be adequately covered here. For our purposes, it is sufficient to note that in some cases it is impossible to obtain a scale that will, for production control purposes, weigh with sufficient accuracy and speed to detect gradual shifts in average within normal control limits. For instance, if the best scale for a given situation can weigh accurately to the nearest hundredth of an ounce, it will be difficult to use this balance to control an operation whose normal total variability is $\pm.02$ hundredths of an ounce. Under such circumstances, a control technique that will magnify the amplitude of the control limits is required. To this end, the expedient of using the total weight of three samples as the control variable was employed.

The control limit calculations were based on the estimates of standard deviations for individuals. The total weight of N individuals is equal to N multiplied by the average weight of the individuals. The standard deviation of the total weights will be N times the standard deviation of the averages. The standard deviations of averages of N samples is equal to the standard deviations of individuals divided by the square root of N. It follows then that the standard deviation of the totals of N samples is equal to the product of the standard deviation of individuals and the square root of N.

In establishing an estimate of the standard deviations of individuals, one is confronted with the problem of obtaining a proper estimate. Actually, there are two valid sources of such an estimate available. The standard deviations can be computed directly from frequency distributions such as shown above. Such an estimate may include elements of variability which are controllable if identified as such.

Another legitimate and very useful source of an estimate of sigma is the residual of the analysis of variance. The residual variance is, in a sense, a measure of the inherent variability of the machine to produce uniform packages making the assumption, of course, that all heads are adjusted to fill essentially the same average weight. In some cases, this statement would have to be modified to include sources of variability which may be inherent to the operation of the machine only because they are uncontrolled. For instance, we occasionally encounter a situation where there will be a significant variability introduced and reflected in the 'between sample' effect which actually reflects the variability due to differences in the product packed at different times. Very often, such variances are uncontrollable or controllable only at great expense. In any case, it should be recognized that the use of control limits, based on estimates of a filling machine's inherent variability, may result in a situation where the control limits appear to be extraordinarily tight. This would, of course, depend on how much of the controllable variability had been eliminated.

In our example, the average residual variance is .78. The corresponding standard deviation is .87 which can be rounded off to 0.9 units. Control limits for individuals will be ± 2.7. If one arbitrarily establishes, as a criteria of acceptable filling, the requirement that not more than 2.5% of the individual package may be underweight, the target average for individuals would be two standard deviations over label weight. The control limits for the total weights of three random samples will be $\pm 3\sqrt{3}\ (0.9) = \pm 4.8$. The process average must be set at the total weight equivalent to two standard deviations of individuals over label weight if the specified requirements are to be met, (i.e.) 3.2 units.

It must be remembered that such a system will permit the control of machine average only and is unable to detect or control variation in individual weights which may develop as a result of slight shifts in averages on two or more heads which are compensating. This is to say that one head may be drifting toward the light weights and another head may be drifting toward the heavy side simultaneously, with the net effect that such variance may not be reflected in the control of machine averages. It is therefore necessary that the average weights issuing from the individual heads be regularly checked. This is done by periodically sampling units filled in each head. The total weights of three

consecutive samples is then determined. It is quite obvious that there will always be some variation in the average weights of the individual heads. However, the only problem that has to be resolved is defining the limits at which individual head adjustments must be made. For this purpose, we have computed warning limits employing two standard deviations of the totals employed for machine average control. These warning limits are useful to the machine operator since two successive weights between the warning and the absolute tolerance limits are considered as sufficient justification for a head adjustment.

The overall control of the line then is accomplished by frequent random sampling of packages filled by the machine. Three of these samples are simply emptied into a suitable container and weighed on an appropriate balance; the weights are recorded automatically and adjustments are made when necessary. Supplementing this basic control, and as a part of regular procedure, three consecutive samples are periodically drawn from each head and the total weights of the three samples are determined to ascertain that all of the heads are filling to substantially the same average weights as defined by the established limits. The frequency of sampling varies from process to process and depends entirely upon the time rate change that the filling machine is capable of exhibiting. Sampling should be frequent to preclude the possibility of sudden shifts in average or variability going undetected by the machine operator.

By way of recapitulation, filling line weight control is often complicated by unknown or unevaluated sources of variability. Simple variance analysis techniques are extremely useful in eliminating obscure sources of variability which would otherwise go undetected or unappreciated. A few such sources, not specifically considered in this report, are temperature, relative humidity, and product density which have a decided influence on the uniformity of filling line performance. Variance analysis techniques readily lend themselves to the investigation and evaluation of such factors.

Specific control techniques have to be tailored to satisfy the particular requirements of each problem. In some cases, like the one discussed, it is necessary to make provisions for limitations of weighing equipment particularly with regard to speed. The proposed technique of employing the total weight of several samples as the controlled variable is useful and has many possibilities. It can, for instance, be used in developing a program of control by gross case weights. This is a subject in itself and involves consideration of package material variances. However, under proper conditions, it can be extremely effective from a control standpoint and very attractive because of reduced sampling and inspection costs. It also has the distinct advantage of adjusting case weights to any prescribed level thereby affording any desired degree of assurance that all cases will be equivalent to the declared label weight.

One final consideration has to do with the costs of installing and maintaining control procedures. Quite obviously, it is possible to develop highly refined control programs, which can be extremely expensive. There is a break-even point in weight control where the cost of improved control is equal to the savings due to decreased over-fill. There is little point in pursuing weight control beyond the break-even point, all other factors being equal. In short, common sense is essential in all aspects of the development of a control program.

"SEMANTICS, SCALES, AND SAVINGS THROUGH S.Q.C."

Lindson P. Anderson and Richard D. Kornblum
Armour and Company

The three factors involved in our title: "Semantics, Scales, and Savings" are all involved and inter-related in this discussion of our experiences with Quality Control and statistical techniques at Armour and Company.

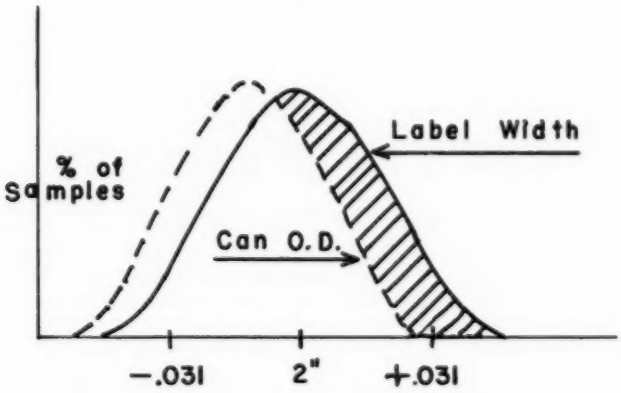
Specific instances dealing with each of the above factors are the most illustrative means of beginning our discussion.

Webster tells us that "Semantics" is the study of the meaning of words. The relationship with S.Q.C. has been well expressed by Lord Kelvin when he said, "that when you can measure what you are speaking about, and express it in numbers, you know something about it. But when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind."

Some of the case histories relating to our semantics problem involve the application of "pure" statistics while others involve more prosaic techniques. We shall start with a statistical problem approach.

Several years ago, we experienced an undue amount of labeling line stoppage due to jamming of the equipment when labeling one specific canned product. Visual inspection showed the labels to be "riding" over the can. The semantic problem then arose - we said the labels were too wide; the label supplier said our cans varied excessively in height. To resolve this semantics hassle, we conducted a statistical analysis of one week's labeling operation for the product involved. We found the following situation existed.

FIG. I



We readily could see that a goodly % of the labels were too large for the can outside width and thus jamming ensued. The label supplier was given these findings through the medium of figures rather than words with genuine cooperation following as to quick improvement in label width variations.

A more routine, but nevertheless very educational problem in semantics began at 8:00 A.M. one Monday when we wheeled a portable electronics "accept or reject" scale into a packaging department that had been producing package weights in considerable assortments, but generally on the high side. We proceeded to route the entire departmental output through the electronic scale, setting the scale so as to reject any packages 1/4 ounce or more over the declared net weight, as well as to reject any light weight packages. By 10:00 A.M. we had accumulated very few underweights but had 50% of the department's production in the overweight reject bins. The outraged foreman erected a semantic barrier by announcing that our electronic scale was "crummy", and we should get a good scale that did not hold up his production. Fortunately, the division superintendent understood the meaning of our exploratory test and, after a short conference with the "cooled off" foreman, initiated scaling reforms among the packaging operators. Later checks have shown a great improvement in scaling excess overweights without creating additional underweights.

Other examples of the use of "clean" statistics to cut through the semantics problem are:

1. Design of, and analysis of results from consumer taste panels. Here we determine whether so-called differences in product could be due to inherently normal variations in taste, smell, and sight, or are the result of control-label factors.

2. Design and analysis of plant experiments so as to meaningfully and economically evaluate new equipment and processes so as to detect true process improvements distinct from normal random fluctuations.

That "economical" reference is an additional asset accruing from an experiment designed by statistical techniques, since these methods of experimental design eliminate excessive testing for unwanted spurious factors, and yet allow for measuring the extent of testing error and interaction between factors under test. Additional benefits realized from statistical experimental design include the randomization factors that eliminate unconscious (or otherwise) biases in judging or evaluating the process or product being tested.

These examples vividly illustrated to us how statistical studies can be of real help in breaking the "semantics barrier" - how they act to stop the confusing and non-productive arguments resulting from lack of reliable data. Statistical techniques are tools in the semantics area - searchlights in any problem that requires measuring and defining.

Let us now turn our attention to the "scaling" part of our experiences.

In the following example, we will refer to the operations of "scaling" or "weighing" as used in our industry. However, the philosophies and facts underlying this example are fundamental to the situation regardless of the specific industry or operation. For "scaling" or "weighing" you can easily substitute the term appropriate to your operation, such as "metering", "measuring", determining chemical concentration, etc.

Virtually every one of our products, and most likely yours, involves weighing (metering or measuring) at some stages of its processing and packaging. The necessity of accurate scales, well maintained by competent mechanics and accurate scaling techniques is vital and even obvious. It is also obvious that all these factors must be adhered to in order to create a profitable operation and yet deliver full measures of quality and quantity to the consumer. It is further obvious that from the accounting, producing, shipping, selling, and buying viewpoints we must have accurate scales and scalers.

But perhaps not so obvious is the extent of the confusion that can result from cumulative weighing errors by scale operators who function at different stages of product processing within a plant.

Let us create a hypothetical example, but one which is familiar to all of us, albeit in different garb for each plant or industry.

For purposes of illustration we shall be concerned with a product "x" (coming from Department "A") and its "shrink". This is the loss in weight inherent in normal evaporation of surface moisture or actual product loss while "x" is being processed for 24 hours in Department "B". Let us further assume that the standard specification for this loss or shrink, agreed upon after much operating experience, is 1.0% or less. Upon this standard shrink is based, among other things, normal operating costs, inventories, and profits.

We then say that product "x" originates in Department "A" and is weighed as it leaves "A" for Department "B", where it is processed for 24 hours and finally weighed, as it leaves Department "B". Thus, shrink % figures for product "x" (which has a standard specified shrink of 1.0%) is the ratio:

FIGURE II

$$\frac{\text{weight at "A"} - \text{weight at "B"}}{\text{weight at "A"}} \quad (100)$$

On the day we are studying the process, let us assume that we have a mysterious insight into the true facts and these true facts are that a 600# batch of product "x" leaves Department "A" and after processing in Department "B" weighs 594#. Thus a loss of 6# in processing gives us the standard specification of 1.0% shrink.

FIGURE III

$$\frac{600\# - 594\#}{600\#} \quad (100) = \frac{6}{600} \quad (100) = 1.0\%$$

Now that we know the true facts, let us go back and see what problems can arise under certain operating conditions.

Our scaling operator at the outgoing scale of Department "A" reads his batch scale slightly to the top of the hairline due to some inherent, but incorrect, bias and reads the original batch weight as 601# (only 1# over the weight we know to be a true 600#). Now the product with its 601# batch weight tag goes to Department "B" and after 24 hours processing is weighed by the scaling operator in "B". This latter scaler reads the after-processing weight as 593#, for he prides himself on reading a little low just to be on the safe side. (This is, after all, but 1# below the 594# true weight that only we know of.)

Now these figures are rushed to our plant accounting department and a shrink figure is calculated.

FIGURE IV

$$\frac{601\# - 593\#}{601\#} \quad (100) = \frac{8}{601} \quad (100) = 1.33\%$$

A shrink figure now has been arrived at that is 33% over our standard specification shrink of 1.0%, which is what we know to be the true correct figure for the day's run. If this situation occurs daily, which it well might, a long chain of events will be set in motion, all based on a shrink figure which is not the true, and correct, shrink value. Some of the immediate consequences of this situation can be:

1. The accounting department must take this 1.33% figure as actual repeated performance, leading to:

2. An expensive time consuming investigation of Department "B" so as to check for defects in equipment, and:

3. A strenuous conference between the plant superintendent and "B" personnel.

But the most serious of all these consequences is:

4. Product cost is increased in paper, which in turn must lead to loss figures for Department "B" and/or an increase in selling price of product "x" which may price it out of the competitive market.

Now, remember that all of the four consequences listed above are based on a succession of two weighing errors - errors that are not dishonest on the part of the scaling operators involved and yet errors that can unleash a series of harmful, and needless, events.

What can we do to assure ourselves a minimum of such scaling variations? We have found the following, rather prominent points, essential in eliminating such situations:

1. Use of the most accurate and durable scales available.

2. Regular preventative maintenance and frequent checks by

well-trained scale specialists available at all times. In our plants a specially trained scale mechanic is assigned to a specific set of plant scales.

3. Adequate numbers of spare scales to act as substitutes while repairs are made on original equipment.

4. Training sessions for plant scalers so as to teach approved techniques, as well as to point out the causes of possible inherent bias. The chain of consequences resulting from inaccurate scale readings should be included in this training.

5. Periodic and unannounced checks of scaling readings by test personnel with immediate correction of any deviations found.

6. Careful attention given the semantics of each problem.

Now why do we concern ourselves with a situation wherein we are so acutely conscious of how close an operator reads the scale to the nearest mark? Let's look at the economics a moment. For example, we'll say this product is selling at 50¢ per pound and we are handling 1500 such lots daily.

FIGURE V

1500 x 50¢ x 601 = $450,000 paid for product

at 1% loss
1500 x 50 x 594 = 445,500 net worth

at 1.33% loss
1500 x 50 x 592 = 444,000 net worth

Net $ 1,500 loss in excess of standard

When a figure such as this is involved, we can readily see the importance of the size of the initial discrepancy that will invariably initiate the chain of consequences touched on previously.

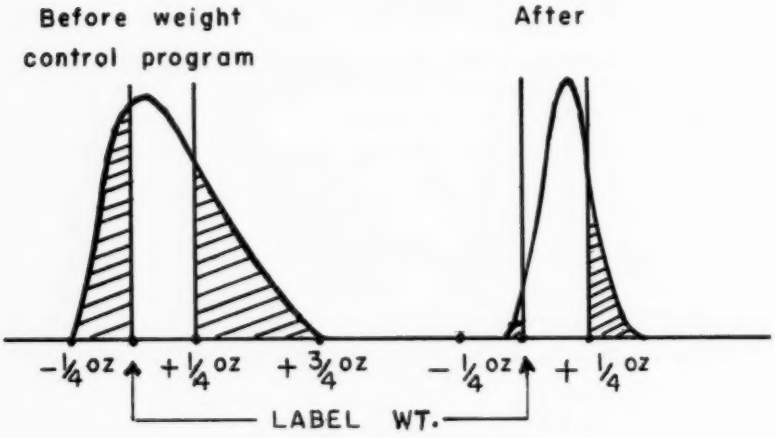
Other specific applications of a quality control nature have involved the factors of semantics and scaling as well as the allied element of savings. In all instances, we have regarded our responsibility as primarily educational, in that the programs after survey and training stages have been turned over to the operating departments concerned. The technical personnel responsible for surveys and plant training were sent by our Research Department to both elementary and advanced S.Q.C. courses at Iowa State and Purdue Universities.

One product survey was primarily a single-plant program aimed at (1) reducing excessive overweights in consumer-size packages while holding underweights to a minimum, and (2) controlling products and package appearance to a desirable standard. To these ends, after training courses were given to plant scalers and inspectors, individual operator control charts were used as a basis for summary sheets to management which pin-pointed operators and/or scale defects, as well

as improvements in scaling and packing.

As a "second-check", a weekly shipping room audit was made by supervisors. Money saved by controlling overweights paid for the inspectors salaries, and we accrued the additional benefits of virtually eliminating underweight complaints as well as acquiring increased sales due to improved and consistent package and product appearance.

FIG. VI



Of a similar nature as to points covered was a program begun on a line of products produced in six major plants that included consumer-sized packages as well as institutional bulk items.

As an example of incoming inspection, we have a sequential sampling plan in regular use on a product container that insures minimum production line stoppages due to faulty "empties", as well as assurance against consumer complaint due to container imperfections.

We have only touched on the far-reaching impact of semantics, "scales", and "savings", in our business as well as yours. Perhaps your industry measures where we scale; but be it calipers, micrometers, metering, or electronic analyzers, the same principles apply in the gauging of process accuracy. Perhaps you refer to an outside diameter tolerance or an upper chemical limit where we discuss length of slice, - you may speak in ten thousands of an inch, we in 1,000# batches, but regardless of the industry, the same immutable principles apply and the consequences, and benefits, extend to dollar savings, consumer satisfaction, repeated and new sales - up to the very continued existence of the concern in todays competitive market.

QUALITY CONTROL FROM BENCH TO BUDGET

Frank H. Squires
Topp Industries, Inc.

Statistical quality control, or more commonly, quality control, is an ascetic among industrial techniques. Calm, aloof, detached, faintly academic, it appears to move through the clatter of the machine shop and the organized confusion of the assembly lines like a visiting professor. Concerned only to bring light into the unanalyzed darkness, it appears incapable of arousing emotion.

And yet it has aroused violent emotions in the shop and in the carpeted offices on mahogany row. It has had the effect of a bull in a china shop without ever feeling like one. It has gone into the big industrial shops like a good samaritan, mounted on an abacus instead of a horse and it has been received with suspicion and occasional resistance.

Why can something so innocent cause so much excitement? Because it isn't as innocent as it appears; it is more than a set of statistical techniques....it is a challenge to the emotions, to the intellect, and to tradition.

It is this challenge on all fronts which makes quality control the tremendously interesting profession it is and which creates the difficulties I want to describe.

On the bench, in the machine shop, quality control engineers seek to install those excellent techniques which appear to be self-evidently good and desirable. But they are resisted and sometimes rejected.

What goes wrong? There has been insufficient preparation, and the failure to prepare is caused by a lack of understanding for the particular problems of the shop and the relationships which exist between inspectors and production operators.

Consider an inspector in the machine shop, a skilled machinist and set-up man promoted to inspection. A quality control engineer is sent down to the shop to install Shewhart control charts. Engineer and inspector meet to discuss the same problem, but they don't speak the same language--they approach the problem from different angles, and in next to no time there is mutual exasperation!

The inspector's entire experience has taught him to see the parts piece by piece, each a distinct and separate identity. As a machinist he once labored over each piece with equal care. He was occasionally distressed by unaccountable variations but he did not doubt that an extra effort on his part would produce identical parts.

The first clash occurs when the quality control engineer turns impatiently from the inspector's careful demonstration of a defect and asks, "how frequently does it occur?" The very phrase has an alien sound amidst the oil and noise of the machine shop. It is the first indication that the engineer does not "see" the individual piece, except as a small and insignificant sample of the group. This becomes more evident as the conversation continues and the engineer presses for the process average quality level in recent lots, or for information from which it might be computed.

The engineer shortly reveals that his aim is to control the process so that the probability of making defectives will not run higher than 1% or 1½% or whatever quality level is established. This may be a far higher yield of good parts than the process has ever given but the quality control engineer's willingness to settle for anything less than perfection shocks the inspector. The inspector and the machinists he works with have always had only one aim: perfection! Although it is never achieved there is great virtue in the effort and an equal virtue in the continued hope and expectation. By contrast it seems almost immoral to abandon the search for perfection and to become reconciled to "second class" goals like 99%, or even 99.9% good. The engineer is familiar with and accepts the concept of the imperfectability of human effort; the inspector isn't, and doesn't, and neither does the vast majority of mankind.

There is the matter of "judgment" or "common sense." An inspector operates under the protection of those powerful and popular labels when he considers the possibility of accepting parts which do not quite meet the specification. He may do this because he believes, from his knowledge of the end product, that the slight discrepancy will do no harm; or because he believes the tolerance on the mating part will compensate for it; or because he has seen similar discrepancies accepted in official engineering review in the past; or because he knows the part to be "the best" the man and the machine can produce. The tendency to award high marks for effort and to take the position that a man's best ought to be good enough is not confined to inspectors. It is another universal human trait and the individual who intends to challenge it must prepare a powerful case. The quality control engineer challenges it without being prepared and is surprised by the resentment that flares up. The engineer is unprepared for what he will find; the inspector is equally unprepared for what is expected of him. Furthermore, the change of procedure, almost 180° out of phase from what he has been doing, is proposed to him by a stranger on his own ground, which is adding insult to injury.

The inspector will be required to comprehend a new procedure, the setting up and maintenance of $\bar{X}$ and R charts. These require some knowledge of mathematics, no matter how simple the working techniques are made. He is going to be required to let the job run when he knows there is a probability it is producing some small fraction of discrepant parts, which challenges his sense of propriety. Most unfortunately he is going to have to surrender the exercise of "judgment" and "common sense" as these are commonly understood in the shop and be guided by careful measurements objectively recorded on the control chart. From having been the all-powerful shop oracle he is to become an attendant for an inanimate control chart, or so it seems to him.

The inspector is confronted with a direct intellectual and emotional challenge. He must learn something relatively complicated and he must surrender jealously guarded power.

The intellectual challenge should not be presented on the floor at all. Inspectors should be thoroughly trained in their use before control charts appear in the shop; it may thus develop that the inspector will become the explainer, the expositor of the new technique and will discover an emotional satisfaction akin to the exercise of power.

There are other solutions for these very serious difficulties but the battle is already half won when the quality control engineer is aware of them and approaches them with sympathy and understanding.

BUDGET

It would be expected that an austere objectivity and a frigid regard for the facts would dominate the negotiations for a quality control budget. Not at all! Here also there are emotional overtones and the occasional clash of conflict between current needs and traditional attitudes.

Every quality control administrator has learned that his most important duty is the preparation of the budget, "selling" it to general management and keeping it sold.

He must be prepared to explain convincingly why the allotment for quality control should be set at a certain figure. He will have to meet frequent proposals to modify it, to rationalize it, to reconsider it, all of these being euphemisms for "reduce it."

The limited enthusiasm for quality control budgets can be explained; there are two strikes against it before the budget conference opens.

Strike one: Quality control is labeled indirect or non-productive in accounting terminology.

Strike two: Quality control was largely identified with inspection and the identification has stuck.

Accounting terminology has enshrined relationships which existed between the various departments of an organization at the beginning of the century

In those by-gone days products were simpler and industrial organizations less complex. In the Victorian period one can imagine the dimly lighted factory, the rows of manual workers bending over benches and machines; the owner in a stovepipe hat stands at the end of the shop, counting the finished products coming off the end of the line. The Wells Fargo Express is waiting, ready to rush them to all parts of the country at breakneck speed.

There were no complications, the workers got their wages and the owner got the tax-free profits.

When the business got bigger, the owner's daughter-in-law came in to keep the books. The manual workers' wages might then have been termed productive, and the daughter-in-law's non-productive. The ratio of non-productive to productive work was obviously a very small one. This tiny ratio became a precious legend.

Times and products have changed Products of incredibly increased complexity and precision are being turned out in quantities beyond the wildest dreams of the stovepipe hat. These products require thousands of hours of preparatory engineering, production and inspection planning, the expenditure of great fortunes in machine tooling in order to reduce the actual manual labor to a minimum. Without these preparations and without their continuous support the raw material would rot and rust and machines and men would stand idle.

The manual labor content has diminished tremendously and the indirect increased accordingly, but the memory of those early days lives on. That almost prehistoric ratio is remembered with anguish when modern "non-productive" departments present their budgets.

We are heading into automation where the required amount of manual labor will shrink almost to zero and still the beautiful legend lives on. Hordes of engineers and programming attendants will wait on the automated machinery and the ratio of "non-producers" to "producers" will become astronomical.

Since the real work will already have been done by the engineers and planners (as it is to a large extent even now), all that will remain for "direct labor" will be to push the button.

It would emphasize dramatically the profound change which is taking place to allow the owner's daughter-in-law to push the button.

This charming person will then have moved from the numerator of the ratio to the denominator and will finally stand the beautiful legend completely on its head.

The second strike against Quality Control is its early identification with inspection.

In the late '20's and early '30's when Walter Shewhart's statistical quality control procedures made their first appearance in industry they were put into the hands of inspectors. The intent of the quality control procedures was to prevent the manufacture of defective parts. Quality control was to be an integral part of the production plan, its particular aim being high productivity and a planned elimination of scrap. But inspection was popularly recognized as a sorting operation which went into effect only after the production job had been finished. In fact, it was said that inspectors neither designed the product, nor made it, nor sold it, but only rejected it.

This was a most unfair appraisal of inspection but it had a popular vogue and, because of the close association, quality control was similarly stigmatised.

Time has modified this attitude somewhat but not entirely. From the bench to the budget the Quality Control Administrator must be sensitive to the emotional climate; he must remember that the traditional view of "indirect" or "non-productive" labor and the effect of the early identification with inspection are inherited from the past and have therefore, in many eyes, a certain validity. They can only be dispelled by constant demonstrations of the productive character of quality control. Sympathetic understanding is the watchword for the quality control engineer whether he is making an installation in the shop or presenting a budget to general management.

ON SOME OF THE NECESSARY INGREDIENTS OF STATISTICAL STANDARDS OF SAMPLING

By

W. Edwards Deming
Graduate School of Business Administration
New York University

WHEN IS THE RESULT OF A SAMPLE ACCEPTABLE? The basis for a statistical standard is statistical theory, good statistical practice, careful execution of the instructions for sampling and for the testing or for the interviewing, agreement on the meaning and limitations of statistical calculations, and usage. Statistical theory and practice have arrived at a stage where there is universal agreement on certain practices and interpretations. For example, any two competent statisticians will agree that certain sample designs, if carried out, will give results that possess valid standard errors. They will agree that other procedures will not do so. Moreover, they will agree on the calculations of the margin of sampling error for a stated probability, and on the amount and kind of knowledge contained in the standard error.

The margin of sampling error of a prescribed sampling procedure refers to the margin of difference, for a stated probability, between the result of applying this sampling procedure and the result of an equal complete coverage of the same frame. An equal complete coverage is a coverage of all the sampling units in the same frame that the sample was drawn from, carried out with the same inspectors or interviewers, with the same instruments and definitions, same procedures, and with the same care as was exercised on the sample.

A good sample picks up samples of good work and of careless work, of correct measurements and of incorrect measurements, of correct responses, incorrect responses, nonresponse, mistakes, all in about the

the same proportions that would appear in an equal complete coverage.

Actually, as is well known, the result of a sample, especially for a study that is difficult, is for many purposes demonstrably more reliable than the result that an actual complete count would give. The reason for the difference, when it exists, lies in the superior workmanship that is possible in a small sample. However, superior workmanship in either a sample or a complete count does not just happen: it requires knowledge of the material, knowledge of how people perform their duties, knowledge of theory, skilful planning, and directed effort in the training and in the supervision throughout the job.

PRECISION DOES NOT GUARANTEE USEFULNESS. High precision (small standard error) in the result of a survey or experiment only means that an equal complete coverage of the same frame would give very nearly the same result as the sample gave. However, the usefulness of a result depends not only on its precision, but even more basically on whether it extracts information that leads to new knowledge, or is a real help in the solution of some problem. Highly precise information that we can not use is little comfort.

High precision does not imply (a) that the frame was complete or relevant to the problem, nor (b) that the method of test, or the questions or the method of interviewing were relevant, nor (c) that the performance on the job was good.

In more detail, we note that the result of any survey or experiment, whether by a complete coverage or by a sample, can only refer to the frame, which is the material presented for study. The frame is not necessarily the universe, which is the material that one really wishes to know about. Unfortunately, in many problems, any frame within reasonable cost will fail to cover a portion of the universe.

It may be important to note that, in a state of statistical control, the frame (material manufactured) may be taken as a random sample of the universe, and that objective inductions may be made by the theory of probability from a sample of the frame to the production process that created the frame, and hence to causes of variation or to causes of the level of quality.

The test of a frame is this: if an equal complete coverage of all the units in the frame would be acceptable scientific or legal evidence, and if it would furnish a useful result, then a sample from the same frame would also be acceptable evidence, and would be useful, provided the precision of the sample is sufficient for the purpose.

This answer does not tell us what frame would furnish useful results, nor what information about this frame would be useful, but it does answer a very important question about sampling. It throws the burden on to the equal complete coverage, and on to the experts in subject-matter, where it belongs.

Confusion between (a) the universe and (b) the frame; and between (c) the precision of a result and (d) the relevance of the information derived from the prescribed method of test or of interview, have been responsible, I believe, for a great deal of misguided counsel in respect to the uses and misuses of data from both samples and complete coverages. This paper aims to clarify some of these points, and to put the responsibilities where they belong.

WHAT ARE THE INGREDIENTS OF A GOOD SAMPLE? Strict adherence to the theory of probability in the sample-design, such as the use of random numbers in the selection, is a necessary ingredient, but this is not enough to ensure useful results. We must build into the survey or experiment, in addition, certain other ingredients, which must come

from (a) knowledge of the subject-matter (chemistry, physical science, production, law, accounting, sociology, marketing, etc.) or from just plain understanding of human nature and of the relationship between product-design and the consumer; (b) careful execution of the sampling plan; and (c) measures of the nonsampling errors, and an evaluation of their possible effects on the accuracy of the results derived from the sample.

The statistician with his theory is helpless if the company that he works for prescribes tests or questions that will yield information of little use, or if the employees of the company are careless in carrying out the work that the sample-design calls for. If a company or a client depends on the statistician and on statistical theory to take over responsibilities that only the company or the client can discharge through knowledge of their own problems and through substantive knowledge, any survey or experiment may too easily turn out to be a disappointment. There are many ways to fail. One way is to fail to achieve as much good information as would have been possible with better planning.

The statistician and the company or client that the statistician serves have well-defined responsibilities, and it is dangerous to exchange them or to get out of bounds. The most important step in the near future toward better statistical work in industry will be better understanding of these responsibilities. I have learned in my own practice that it is important to tell the client at the outset what are his responsibilities to the job, and what mine are. A code of professional statistical service may be based on what follows.

The reason for stating these responsibilities explicitly here is (a) to enhance the usefulness of statistical theory and practice; (b)

to forestall disappointment on the part of the client, who if he fails to exercise his responsibilities in the planning of the survey or experiment, may not realize in the end its fullest possibilities, or may discover too late that certain uses that he intended to make of the results are impossible.

RESPONSIBILITIES OF THE STATISTICIAN. I shall in what follows use the word client for simplicity, but what I have to say applies equally well to the relationship between a statistician on a salary and the company that he works for.

The statistician's responsibility in any survey or experiment is only for its statistical aspects; specifically:

a. to design a sample of a specified material, covered by a specified frame, to reach a desired degree of precision; or to design an experiment, to reach significance in the tests of specified materials or methods or types;

b. to explain to the supervisor in charge of the work the instructions for the selection of the sample and for the computations; to satisfy himself that the supervisor understands the instructions and the possible effects of departures therefrom. The statistician should make it clear to the client at the outset that he has the privilege of withdrawing from the engagement at any time if he feels not satisfied with the client's performance on the job.

c. to design a plan for probes of the execution of the work, to detect and to measure the extent of the possible nonsampling errors of chief importance (listed in some detail later on under the heading "The statistician's report or testimony");

d. to construct upon request of the client statistical aids to supervision, to improve the uniformity of the tests or of the interviews;

e. to explain to the client, when the tabulations are finished, the meaning of the results of the survey in terms of their statistical significance. However, the statistician will not recommend that the client adopt any specific administrative action or policy. The uses of the data obtained by a survey or experiment are in the end entirely up to the client.

The sample-design will define the sampling units. It will explain how to classify and number them, and in some designs, how to thin one or more of the strata. It will contain instructions on the computation of the estimates desired, and of their standard errors. The selection of the sampling units for the sample will at every stage be made with random numbers. The statistician will furnish these numbers, or will prescribe definite rules for the use of a table of random numbers, only after the client certifies that the preparation of the frame, including the scheme for giving a serial number to every sampling unit, is complete.

The design of a sample or of an experiment calls for skilled use of statistical theory, coupled with knowledge of men and of materials. Judgment and knowledge of the subject-matter, so vital in their place, can not design samples nor experiments, nor select the sampling units for test. The statistician must be firm on this point.

RESPONSIBILITY OF THE CLIENT. The client should state in advance how he expects to use the results of the survey or experiment. Otherwise, no proper statistical design of a survey or experiment is possible. It is the responsibility of the client to decide whether the information that he needs can be elicited at all from a sampling unit in the frame. If not, no survey, complete coverage or sample, will be satisfactory.

The development of the plans for a survey or experiment may require experimentation and trial, with successive revisions. The client will carry out the statistician's instructions for these explorations, and will make the calculations therefor, with no changes in procedure without authorization.

The plans, when finally fixed, should be in writing, and the client will make no changes in procedure without authorization so long as the statistician's responsibility remains in force. This places a responsibility on the statistician: he is subject to call night or day, and he may, in his judgment, need to make frequent enquiries about the work.

The client should arrange for the statistician to have direct access at any time to the people that carry out the preparation of the sample, the testing or the interviewing, the supervision, and the computations.

It is essential that the client know in advance what portion or classes of the universe any proposed frame covers, and how thoroughly, and what it does not cover. The responsibility is the client's to decide whether a proposed frame and supplementation if any are sufficiently complete for his purpose. The test is whether a complete coverage of the frame would suffice.

The client and not the statistician must assume the responsibility for those aspects of the problem that are substantive. Specifically, the client must:

a. decide the type of information that the survey or the experiment is to elicit

b. prescribe the methods of test, examination, questionnaire, or interview by which to elicit the information

The client must assume responsibility for:

c. the decision on whether the frame is sufficiently complete

d. the statement of the classes and areas of tabulation

e. the actual work of preparation, training, testing, interviewing

f. the supervision of such work

g. the completeness of coverage

h. the coding

i. the tabulations and the computations

The statistician may of course work with the client in these matters, but the ultimate responsibility for settling them is the client's.

THE STATISTICIAN'S REPORT OR TESTIMONY. The statistician's report or testimony concerning a survey or experiment will include the estimated standard errors of the results of chief importance, the calculations of significance, and their statistical interpretation. It will include also information derived from any probe that he requested for measurement of the nonsampling errors, such as:

a. failures to select the sampling units in the manner prescribed

b. failures to reach and to cover sampling units that were selected or should have been selected

c. the inclusion of sampling units not intended for the sample but covered or partly covered by mistake

d. other slips and departures from the prescribed sampling procedure

e. errors and difficulties in reporting

f. nonresponse

These errors are measurable by observations on a properly selected subsample of the main sample. The statistician's report or testimony should include an evaluation of the possible effects of any such blemishes on the conclusions drawn or to be drawn from the data.

It is interesting to note that measurement of the nonsampling errors is as important in a complete test or census as it is in a sample. In a complete census, the nonsampling errors are measured by a sample of a few thousand small segments of area, re-covered by expert interviewers. Verification procedures by a sample often turn up amazing discrepancies and omissions in so-called 100 per cent tests and audits.

The statistician's report will not urge any specific decision upon the management of the company on the basis of the results of the statistical study. This responsibility belongs to management, which may be in the possession of information not provided by the survey. Moreover, no decision can have the objectivity of a statistical result. To urge a certain decision on the basis of the results would be to cloud the results of the survey.

The statistician's responsibility carries with it any writing that may appear in print with his name attached in any way as a participant. There should be an understanding that the client, if he prints or publishes the statistician's report, will print it in full, and will not omit any part of it without the consent of the statistician.

There should be no necessity, with the adoption of statistical standards, to present in legal evidence the theory of sampling, nor the details of the sample-design that was used, provided it was a probability design. The standard error, and the number of degrees of freedom in the estimate thereof, plus some relevant measure of any serious departure of the distribution of the estimate from normal convey all the information

that there is to convey in respect to the margin of difference, for a prescribed probability, between the sample and an equal complete coverage. Knowledge about the details of the sampling plan adds no new information about the sampling error in a result.

Measures of the nonsampling errors add a great deal of information about the accuracy of a result. We may thus conclude that the only relevant evidence about the accuracy of a result is a statement by a competent statistician (a) that the sampling plan was in fact (or was not) a probability sample of a certain frame; (b) that the margin of error for a prescribed probability (e. g. 1% in the upper tail) has a certain magnitude; (c) that probes of the execution of the sampling plan revealed certain departures or no departures from the plan prescribed; (d) that in his opinion these departures if any will likely cause certain inaccuracies in the results, which he will evaluate and describe.

A statistician's certificate could well be a paragraph that describes the sampling plan simply as a probability sample drawn up according to standard and accepted theory and procedures, followed by notations of exceptions and their possible effects. The sampling plan and the frame or frames, and all pertinent details, should of course be available for probe if any question should arise about the execution of the plan, or about any part of the plan itself.

MANAGERIAL STATISTICS APPLIED TO THE BREWING INDUSTRY

Alex G. Nagy
Lone Star Brewing Company

I. General

The importance of the management function in business administration has grown tremendously during the twentieth century. This development, which has been particularly phenomenal during the past generation, has come about as the result of the gigantic growth of industry. The size and complexity of business organizations has necessitated an increasing amount of attention to formulation of plans, budgets, and administrative control operations. The owner-operator of a small business could possess firsthand knowledge of details of his entire operation. This is not so with businesses owned by numerous stockholders, guided by boards of directors and managed by staff executives, none of whom can become intimately acquainted with the multiplicity of details concerning sales, production, inventories, finance, personnel, and administration.

In discharging their managerial functions business executives in all departments have become increasingly dependent upon accounting and statistics.

Managerial statistics deals with data and methods which are useful to management in planning and controlling organizational activities.

PLANNING is the first step in scientific management. All plans are based on forecasts. Formally or informally - knowingly or unknowingly - businessmen are making forecasts when they purchase raw materials, when they purchase plants and equipment, when they engage personnel, in fact, whenever they perform any act in anticipation of the future. Businessmen cannot avoid making forecasts, they can only neglect to make as intelligent a forecast as possible. True, the most certain thing about the future is its unpredictability. However, some aspects of the future can be predicted more reliably than others. The wise businessman reduces the likelihood of error in his forecasts by carefully estimating the future impact of those factors which can be forecast with the greatest degree of accuracy.

ADMINISTRATIVE CONTROL. Well-formulated plans are of little value if they are not effectively carried out. It behooves the business executive to exercise administrative control of his operations in an effort to correct controllable deviations from planned operations. For example, the trend of sales by products, territories, and salesmen is highly significant information. Executives must act rapidly to capitalize on favorable sales trends and also to adjust to unfavorable trends. Similarly, production costs should be subjected to critical analysis. Lack of control of production costs can easily result in a loss rather than a gain from business operations. Reduced costs increase unit profits just as surely as do increased selling prices.

II. Use in the Brewing Industry

Table 1 compiles some of the aspects where statistics have been used successfully. All these activities are part of or related to forecast and internal control, the two major fields of application of managerial

WHERE MANAGERIAL STATISTICS HAS BEEN APPLIED

In The Area of	Managerial Statistics Helped To Determine		
Management	Capacity - Volume-Profit Ratio	Trend Analysis	Administrative Control
Production	Production Schedule	Quality Control	Production Cost Control
Sales	Sales Forecast	Market Analysis	Sales Budget Control
Advertising	Adv. Expense Breakdown by Media, By Sales Districts	P.O.P. Material Movement	Adv. Budget Control
Public Relations	Distribution of Expense by Sales Districts	Use of Special Equipment	Company Publications
Purchasing	Sampling of Incoming Materials	Acceptance Standards	Vendor Rating
Finance	Gross Income Forecast	Performance Level Evaluation	Stockholder's Report
Personnel	Classification and Analysis of Data	Injury Rate vs Loss Time	Labor Negotiation
Traffic	Automotive Cost Analysis	Efficient Routing	Preparation of Data on Rate Hearings

TABLE 1

statistics.

A. Forecasting

To aid management in planning, data to be collected, to be tabulated, to be analyzed and evaluated may be summarized briefly as follows:

1. Factors affecting beer demand:

   A. Population

      -population trends
      -age group development
      -wet and dry issue
          -population by sales districts
      -characteristics of population
          -number of farm homes
          -number of people engaged in manufacturing
          -industrial development
          -Latin-American labor migration
      -race
          -geographical distribution of
              -White
              -Latin
              -Colored population

   B. Disposable Income

      -average effective buying power
          -by sales districts
              -per capita
              -per family
      -cost of living index development
      -economic development

   C. Price of Beer
      -distributor prices to retailer
          -by packages
          -popular vs premiun priced beer
      -profit structure comparison
          -on distributor level
          -on retail level
      -effect of taxation on price, and related to per capita consumption

2. Inter-industry Competition

   A. Economic Importance of the Brewing Industry

   B. Effect of Competitive Industries
      -liquor
      -wine
      -soft drink

   C. Prohibition Sentiment

3. Intra-industry Competition

A. The Whole Brewing Industry

- -beer sales' trends
- -package movement
- -seasonal indices
- -number of breweries
- -methods of acquiring more brewing capacities
- -development in rank of sales
- -the place of a regional brewery in the industry
  - -local advantages
  - -purchasing disadvantages

B. Local Beer Sales

- -factors affecting local beer demand
- -local beer sales' trend
- -out-of-state vs local brands
- -package movement
  - -effect of dry territories on package movement
  - -pricing effect on packages (popular vs premium)
- -brand comparison analysis
  - -by brand, by package
  - -by geographical distribution
  - -effect of beer sales reporting regulations on the reliability of analysis
- -market evaluation
  - -geographical distribution of market potential
  - -buying motives
    - -brand preference
    - -package preference

The above information, supplemented by the brewery's own operational data then becomes the basis for the sales forecast, regardless of whether management is using an intuitive or a mechanistic approach. There are many books published on forecasting, and they suggest a number of alternative methods of procedure. The most widely used methods are:

(1) correlation and regression analysis
(2) curve fitting or time series analysis.

Many companies place substantial reliance on sales forecasts by salesmen and by surveys; thus using

(3) salesmen's opinion
(4) consumer surveys

The newest methods combine the wisdom of experienced businessmen with statistical analysis in the

(5) filter technique
(6) skeptic's technique

Let me summarize the advantages of a sound forecast:

a. Successful budgeting of expenses, costs and profits depends on good forecasting of sales income.

b. Successful forecasting reduces the area of avoidable risk. Forecasting will seldom, if ever, be without some error; but forecasting can limit the area in which guesswork is the only guide.

c. Good forecasting can stabilize production and employment over the years by ironing out variations caused by seasonal fluctuations of sales. Steady employment can mean better labor and community relations, lower employee turnover and lower labor costs.

d. Better forecasting will be needed by management to deal successfully with the growing rigidities of labor costs and other problems brought about by the demand of organized labor and public opinion.

e. Satisfactory control of inventory of all kinds -- brewing raw materials, packaging materials, beer in aging and finish tanks -- is dependent on satisfactory forecasts of future sales. Successful planning of long-term investment programs and of corresponding new capital requirements depends on reasonably accurate long-term forecasting of sales.

f. The successful use of standard cost systems for cost and expense control, and for satisfactory pricing of products depends on good long-term forecasting of sales and production volume.

B. Managerial Control

The making of decisions is in a sense both the first and the last step in a continuing, never-ending, dynamic managerial control process. Ideas for the conduct of a business are originated, evaluated, decided upon, and put into effect. The resulting action is then appraised in terms of the objective to be attained, expressed as greater revenue, as lower cost, or otherwise. This appraisal gives rise to a new proposal, an adjusted program. Further plans and decisions are adapted to new needs and changed external circumstances.

The need for control, in this sense, is of comparatively recent development. This need grows directly out of the delegation of authority and responsibility characteristic of modern, large-scale enterprise. In small businesses control is effected by direct observation and supervision, and by the personal knowledge of the owner or manager and his contact with each individual transaction and aspect of operations. This knowledge and contact of top executives is not possible in a large organization, although it is present, to a degree, in the lower ranks of management. Top executives must nevertheless supervise the activities within their respective jurisdictions despite lack of personal and direct contact with events and personnel. This supervision is accomplished by the delegation of authority and responsibility to lesser executives and by the introduction of other means of managerial control in substitution for the direct knowledge, contact, and action that are no longer possible. The new agencies of control that are now necessarily introduced are the carefully formulated and predetermined objectives, policies, plans, programs, standards, and cost system. These instruments are not mutually exclusive and independent; rather, they are interrelated phases of mana-

gerial effort to secure and maintain control under conditions of delegated authority.

Control, as the word is used above, may be applied at all levels of the organizational structure and to all phases of operations.

All too often accounting records are maintained solely to provide data required for payrolls, customers' accounts, and for reports to stockholders, creditors, various regulatory bodies, and the Internal Revenue Service.

A brewery's accounting records should be geared to its organizational structure in order that data drawn from the accounts may readily reflect the degree to which executives and supervisors have discharged their responsibilities. Accounting systems designed on a responsibility basis yield data required by nonmanagerial groups, and in addition, facilitate preparation of reports required for internal administrative control.

Accounting and statistics have been long recognized as primary tools of the business administrator. To be sure, accounting is much more firmly established than is business statistics. Nevertheless, increasing recognition has been given to the importance of statistics, and the relationship between accounting and statistics has become more evident. This was noticeable first in the development of cost accounting and more recently in the shift toward responsibility or managerial accounting, in which accounts are used to provide data to measure executive or departmental accomplishments in relation to some desired standard or level.

Business administrators require a variety of reports which often can be obtained only from several classifications of the same data. For example, management often desires reports on sales classified by package, territories, salesmen, and distributors. The cost of maintaining a set of accounts which would provide directly all such classifications would be prohibitive. Rather, primary classification of accounts should be on the basis of responsibility; required secondary classifications may be obtained easily by the use of electric or manual sorting device.

Analysis of trends of profits, sales, and expenses by various pertinent classifications may be described as a function of either the accountant or the statistician. The important point is that whoever makes such an analysis should be well grounded in both accounting and managerial statistics. The area where statistical analysis can substantially extend the scope and penetration of accountants' services is in the field of aid to management.

C. Application of Statistical Control Techniques

BUDGETARY CONTROL: Forecasts of sales - formal or informal - serve as the foundation for plans of future business operations. Integration of all phases of a brewery's operations to the sales forecast may be accomplished by budgeting. A budget may be thought of as a plan for future operations expressed in financial terms; actually, this requires that plans first be expressed in physical terms, such as units of sales, production, and number of personnel. A complete and integrated budget plan would probably include the following budgets:

Sales budget

Inventory budget
Production budget
Purchasing budget
Various expense budget
Capital expenditure budget
Financial budgets

The close relationship existing among the above budgets is quite obvious. However, budgets to be effective, must be realistic. They must also be subject to more or less constant review. When the occurrence of unpredicted events changes forecasts, corresponding revisions must be made in the budgets. The unpredictability of the future does not destroy the value of budgeting. On the contrary, budgeting familiarizes executives with the relationships among the several phases of a brewery's operations and thereby enables them to adjust operations more readily to meet changed conditions.

All budgets may be reduced to control charts, thus providing department heads with a simple line graph, enabling them not only to follow developments by the month but also by comparison to any past period.

From a managerial point of view there is one very important aspect that I would like to emphasize. Accounting records may reveal the person responsible for any mistake or losses; but the mistake need not have been made if he had been cautioned in time. Setting up pre-determined control limits on charts advises department heads of this caution before their action goes out of control and thus will eliminate mistakes and losses. Control limits could be determined by the same techniques used in quality control, or by management decision.

SALES CONTROL: The purpose of sales control is to determine the difference between actual sales and budgeted sales at any given period in order that operational plans may be adjusted accordingly. Since gross income and consequently expenses are budgeted according to sales, performance level analyses are of utmost importantance. Package movements, geographical distribution of potentials, difference of sales between sales to distributors and distributors' sales to retailers should be continuously evaluated.

INVENTORY CONTROL: Inventories of stock should be adequate to meet sales but should not be excessive. Lack of inventories means loss of sales, but excessive inventories mean not only that capital is "frozen" or idle and not earning any return, it could result in some old beer either at the distributor's warehouse or in a retail outlet.

Control charts for each package may be set up indicating in units the sales to distributors and their sales to retailers. These charts will show two important features:

1. seasonal movement differences
2. inventory movements in the territory

Sales, inventories, production, and purchases must be geared together if maximum efficiency of operation is to be realized.

PRODUCTION CONTROL: Brewery's production control is twofold: qualitative and quantitative.

The production director or brewmaster is responsible for the uniformity of the beer produced, packaged and distributed. The brewing quality control program was the first department in the brewing industry to make use of statistical techniques in brewhouse as well as bottle shop control. Some breweries have extensively used statistics in vendor rating and taste panel evaluation.

The brewmaster's efficiency is judged not solely on quality of the beer produced, but also on how economically he can produce it. Efficiency could be achieved by:

1. producing as much as can be sold
2. controlling manufacturing costs

Both of these factors may be set up on charts. A continuous five year chart indicating production, forecast, and actual sales by month, by package; supplemented with a production-sales-inventory chart should give adequate information for short-run adjustments. Significant from a cost point-of-view are the bars representing over and under production. Coordination of production to actual sales and vice-versa achieve the most economic and efficient operation.

Manufacturing cost may be set up on charts by using unit cost as follows:

For Brewmaster: Manufacturing cost

-brewhouse and cellars
-bottleshop
-power department
-overhead

For Brewhouse: Total cost

-raw materials
-labor
-power
-overhead

For Bottleshop: Total cost

-packaging materials
-labor
-power
-overhead

On these basic charts all uncontrollable changes, such as increase in price of raw material, new labor contract, etc. may be indicated. This technique may be applied to any phase of operation which need to be analyzed further, such as composition of raw material and packaging material costs, cost of different packages and lines.

Comparison of manpower production figures with those of similar size breweries, both at brewhouse and bottleshop level, could foster efficiency.

EXPENSE CONTROL: Expense budgets should be supplemented by the use of expense and performance standards for as many expense items as possible.

Actual expenses and performances should be compared regularly with the expense budgets and standards. This enables executives to take remedial action whenever significant differences occur.

Expenses, such as shipping and delivery, selling, advertising and administrative, predetermined by the sales forecast, may be set up on control charts on a $/bbl. or unit sold basis. Using a semi-log paper, the slant of the line will indicate the true changes month by month. Management may determine control limits for the total expense by department which also could be carried on the same chart.

FINANCIAL CONTROL: A thorough analysis of sales and expenses will enable a brewery to construct a break-even chart. Such a chart illustrates the profit or loss which will result at various volumes of sales over a period of time, provided expenses are controlled according to plans, and provided significant fluctuations do not occur in prices of material, labor, supplies, and finished goods. Incidentally, this same technique could be used for determining the distributors break-even point.

Other charts used in financial control:

-geographical distribution of income vs expenses
-effect of taxation on income dollar
-analysis of stockholders by size of holding, sex, geographical locations

REPORTS CONTROL: Internal administrative control is largely dependent upon preparation and analysis of reports on all phases of operations. However, a burdensome reporting system can easily defeat the purpose for which it was intended, namely, effective managerial control. Data should be reported only if they serve a specific control purpose; data should not be collected merely because they are "interesting" or because they might "come in handy sometime." Remember:

1. Measure only what is measurable
2. Measure only what is important
3. Keep it simple and accurate
4. Use it as a tool not a substitute for judgment.

III. Conclusions

Statistical techniques are almost indistinguishably interwoven with accounting and reporting techniques in the area of managerial control. Statistical techniques which have useful applications in control include measures of central tendency, measures of variation, ratios, percentages, index numbers, trend analysis, correlation and sampling. Although managerial control involves much more than the application of statistical methods, it is difficult to conceive of an area of control where some statistical technique could not be employed to advantage.

Administrative and statistical controls, properly designed, highlight essential information which otherwise would remain forever obscured in the mass of detail which necessarily characterizes accounting records.

REFERENCES

Gardner, F.V. Profit Management and Control, McGraw-Hill Book

Company, Inc., New York, 1955.

Hanson, K.O. Managerial Statistics, Prentice-Hall, Inc., New York, 1953.

Lasser, J.K. Business Management Handbook, McGraw-Hill Book Company, Inc., New York, 1952

Spurr, W.A., Kellogg, L.S., and Smith, J.H. Business and Economic Statistics, Richard D. Irwin, Inc., Homewood, Illinois, 1954

_____, Brewing Industry Survey, 15th Ed., Research Company of America, New York, 1956.

_____, Brewers Almanac, United States Brewers Foundation, New York, 1956.

AN APPROACH TO THE VENDOR-VENDEE RELATIONS PROBLEM

Jack Mannion
Northrop Aircraft, Incorporated

The Consumer-Supplier Relationship - Introduction

The operating principle of our vast industrial complex places nearly every organization in the field of Buyer-Seller Relationships, for nearly all must buy and all must certainly sell.

Individual organizations have expended considerable effort to reduce the principles of the buyer-seller relationship to a regularized system subject to equitable control. There has not been, however, an industry-wide effort to provide flexible solutions to the constantly recurring problems of both buyer and seller.

Some of the past work has had remarkably useful results and has provided a base of experience upon which more ambitious structures may be built. For example, experience has proven the value of Vendor Certification and its prerequisits of an acceptable quality control program in the Vendor's plant, the necessity that a vendor practice statistical control methods, and the requirement that the vendor enter willingly into a certification agreement.

Most such certification plans also include provision for the classification of characteristics, the standardization of inspection methods, and the reduction of inspection severity based on acceptable quality history.

All of these factors have proven valuable and most have been extensively documented so that no further comment need be made here.

Another kind of helpful effort is currently being expended in a cooperative effort between ASQC Electronics Division and RETMA in popularized explanation of the meanings and the application of Quality Control techniques to the buyer-seller relationship.

One of the most recent developments in the Vendor-Vendee relations field, and the one essentially responsible for the writing of this paper is the formation of the ASQC Vendor-Vendee Relations Committee. This Committee was permanently organized in 1956 under the co-chairmanship of E. C. Bennett of Northrop Aircraft, Incorporated and W. R. Pabst of Naval Ordinance Proving Ground and consists of about 20 members equally divided between Industry and the Military.

The Committee operates as an organized body for the direction of task force teams assigned to the definition of particular facets of the Producer-Consumer relationship. This paper presents a plan which was originally developed to provide better quality assurance and the achievement of measurement compatibility, but grew to the point where we believe it can help solve a number of the problems faced by both Producer and Consumer.

Some Problems of the Producer-Consumer Relationship

First, let us examine some of our problems from the viewpoint of the

consumer.

What, for example, shall be the basis for our decision with respect to the severity of receiving inspection control? In the airframe industry, as in others, we can classify the vendor's product in accordance with our ability to verify its quality.

First, and least difficult to control, are those products which exhibit fully verifiable quality characteristics. An example of such a product might be an electrical stand-off insulator. The insulator must have a certain length, width, and depth; it must have a resistance such that it will not conduct current when a given high voltage is applied across it. The acceptability of this part can be fully verified provided we can simulate its use environment during receiving test or can calculate the part's behavior under its use environment.

Second, and the most difficult to handle, are those vendor products whose quality characteristics cannot be fully verified short of destructive testing, if at all. This category would include sealed and/or functional items and some chemically processed parts.

We have, for example, several electronic packages which we purchase that require a functional test in Receiving Inspection. Due to the fact that they are designed and built by a vendor and then sold to us as a "black box", no real analysis of the assembly short of environmental testing is possible during receiving inspection except the measurement of the outputs when the specified inputs are applied.

We also purchase printed circuits the conductors of which are made up of copper, nickel, and rhodium - stacked up in that order. If the rhodium plating is done properly, we have no way of knowing that the nickel plating has been applied properly unless we destroy the conductor.

Actually there are a lot of things in this receiving inspection business which you must "take on faith", and the trend should be toward the elimination of some of these "articles of faith" by improving the communication between the user and the producer.

However, even when we can determine the quality status of the vendor's product, our ability to obtain correction of unsatisfactory conditions is, in part, dependent upon the level of control we are able to exercise on the various sources of product.

Most amenable to control is our own, the consumer's, proprietary product. For such product we control the specification and we may usually select the producer, but the problem lies in effective communication and the cost of adequate control on the part of the vendor.

Second, with respect to controllability, is the Military Standard or Industry Standard part. Here we do not control the specification directly, but we help make it. We do have, however, a rather large number of vendors to select from which is the real source of our control. Again the pitfall is the specification interpretation, but here also enters the economic problem faced by the vendor who provides better conformance through increased control effort and thus runs the risk of pricing himself out of business.

The next category includes the producer's proprietary product. We have no control of the specification and the producer can usually sell all he makes. The problem here is one of communication - we to understand what the product will do - he to understand what it must do. Purchased quantities have an influence here.

Last and least controllable, from the viewpoint of the consumer, is Government Furnished Equipment. Here long channels of communication and a generalized specification work against effective corrective action. On the other hand, unsatisfactory product, when issued to a number of contractors, has the singular effect of widespread pressure and sometimes results in effective corrective action.

In the past, industry has generally employed a process of "after the fact" correction to its vendor problem. It has made extensive use of liaison methods in an effort to achieve an acceptable mutual understanding of requirements. It has performed selective inspection and often simply stretched the acceptance criteria when the product was badly needed. But usually it has just rejected the product over and over again.

Recently, however, there is evidence of some really enlightened thinking in the direction of general solutions for many of the Vendor-Vendee relations problems. Certainly Industry is better analysing its product requirements and, with the help of the Vendors, better specifications are being written. More cooperative relations with Vendors are being actively and logically pursued, and the producer's performance is being more accurately and objectively evaluated.

We, as the consumer, must extend these efforts and we must provide for a more clearly defined interrelation of the facets of our relationship with our suppliers. We must extend and still further define our requirements. We must establish better organized methods of indoctrination and communication between ourselves and our vendors. We must learn to help the vendor distribute his control effort in accordance with our requirements for his product. We must rate all of our suppliers fairly and accurately, and we must pay special attention to the establishment of compatibility of measurement between ourselves and our suppliers.

We believe the plan devised at Northrop will provide a solution to many of the problems we have discussed.

The Plan

The plan is comparatively simple in operation but rather broad in scope and flexibility of application.

- It provides for unusually efficient use of variables sampling.
- It achieves a maximum degree of possible control over the various product types previously mentioned.
- It provides for assurance of measurement and inspection method compatibility.
- It makes possible both short and long term correlation of discrepancy and performance data.

It rates vendor performance fairly, accurately, and automatically.

Here then, briefly, is how the plan operates:

The first step is the full analysis of the product we intend to receive paying particular attention to classification of characteristics, acceptance criteria, and qualification requirements.

The establishment of characteristic classification is based on the following considerations, each of which is classified Critical, Major or Minor and from which a summary classification is determined - also Critical, Major or Minor.

- The importance of the vendor's product to our end item. (i.e. What will be the result of failure to the safety or performance of the end item?)

- The importance of the individual characteristic to the vendor's product. (i.e. Will discrepancy or failure of the characteristic cause the vendor's product to fail?)

- The probable reliability of the manufacturing process for the individual characteristic. (i.e. What is the probability that the vendor's manufacturing process will produce an error?)

- The risk factor of error detection subsequent to receiving inspection. (i.e. Are we likely to find an error before next assembly or inspection?)

- The economic consequence of an error escaping detection during receiving inspection. (i.e. Is the part buried in the airframe or scattered in stock bins all over the shop?)

The acceptance criteria, the AQL, the inspection methods, and the Inspection tooling requirements for each characteristic are predetermined in accordance with measurement requirements and a summary classification of each characteristic.

The qualification test, first article inspection and/or pilot run inspection requirements, over and above normal acceptance procedure, for significant quantities of product are predetermined to insure acceptability for characteristics suspected of being marginal.

The vendor's contract is then arrived at by negotiation to insure that full understanding of classification and acceptance criteria is achieved. In addition, the vendor is apprised of other requirements as necessary. For example, the vendor may be required to submit certain variables data with respect to characteristics classified "critical", or we may stipulate specific requirements and methods of accomplishment of corrective action and the establishment of measurement compatibility.

When the classifications have been accomplished and the contract negotiated, the inspection procedure is determined. The resulting inspection plan is described by means of a form called a Quality Evidence Record, one of which is used for each part type. The form is

the inspection instruction sheet and is filed in the appropriate receiving area. Prior to receipt of product, the form will be filled out with the following information:

- Identification of the vendor and part.

- Classification of the characteristics and their associated AQL's.

- The inspection tooling requirements.

- A merit value for each part which is merely a count of the part characteristics to be inspected regardless of classification. The merit value is constant regardless of lot size since errors on a few pieces in a lot constitute a greater risk to the consumer than errors on many pieces. There is also the factor of error propagation in that a single discrepant condition tends to permeate an entire lot. The merit value also measures, to some extent, the consumer's opportunity of detecting a discrepancy.

The Quality Evidence Record has the additional virtues of providing a guide for the inspection and a visual indication of historical behavior for the part he is inspecting. The format is such as to indicate at a glance any repetitively discrepant characteristics.

The received parts are then inspected to the established plan and the inspection results for the part number are recorded on the Q.E.R. under the corresponding characteristics. (See Figure 1)

When the vendor has submitted certification data on critical characteristics (see Figure 2), variables sampling or 100% inspection may be employed, and the results of the inspection compared to the vendor's data. This provision considerably increases the assurance to be achieved from variables data, both from a product acceptability viewpoint and from an evaluation of measurement compatibility.

Accepted parts and lots are designated, for tabulating purposes, by the simple addition of the part's merit value to the standard Receiving Report form which also contains the usual identification and quantity data. (See Figure 3)

In the event of rejection of one or more parts, another standard form is utilized, the Receiving Rejection Report. (See Figure 4) This form normally lists necessary identity and quantity data but displays, in addition, the merit value for rejected parts and the specifics of the rejection, such as:

- The kind of rejection is reported - such as visual, functional, dimensional, etc.

- The letter classification of the discrepant characteristic. This classification has been predetermined as previously mentioned.

- A numerical seriousness index associated with the letter classification of the characteristic. The numerical codes are

applied by the receiving inspector and are determined in accordance with the following definitions:

Code 1 - A discrepant condition which with reasonable certainty will cause the inspection characteristic to fail.

Code 2 - A discrepant condition which does not cause immediate failure of the inspection characteristic, but which, if not corrected, would develop into a Code 1 condition.

Code 3 - A discrepant condition which will probably not cause failure of the inspection characteristic at present or in the future.

Each discrepancy type will receive a demerit rating, depending on the combined letter and number code, as indicated in Figure 5. Vendor reported errors will receive 1/2 the appropriate demerit rate because of the reduced risk inherent in preknowledge of a discrepancy.

When the receiving inspection function is completed, parts are dispositioned and handled through normal channels, and copies of the receiving forms which have been generated are forwarded to tabulating. Tabulating keypunches the information and provides a tape which catalogues automatically all transcribed and computed information concerning the performance of each vendor on a lot by lot basis. (See Figure 6) The tape includes:

- Complete identification of the vendor and the part.
- Complete recording of the rejection data including kind, classification, quantity, and whether vendor reported.
- The merit/demerit performance ratio for each lot.
- The number of lots received and the lot rejection rate for the time period.
- A summary control rating, determined from the vendor's lot and piece performance. This is the control figure upon which is based the vendor review decision.

Once the data have been keypunched, the tabulating function is automatic. It permits sorting by any item or combination of punched data as well as an accumulation of information for any time period desired. (See Figure 7). It also permits, through the use of effectivities, determination of the effectiveness of corrective action.

The above described system is really quite simple and culminates in a useable ratio for the determination of a vendor's performance. The summary rating is the device which initiates vendor review action when necessary and functions as such except in the case of an A1, A2, or B1 coded discrepancy which require immediate action regardless of the vendor's summary rating.

Conclusion

In summary, some of the pertinent features of the plan might bear

further mention:

The use of variables sampling is especially valuable in that it permits comparison of the vendor determined process average and range with those determined by ourselves, as well as the detection of marginal design characteristics, and variation due to time, storage and handling. This feature provides for simultaneous analysis of the compatibility between the measurements resulting from the respective inspection systems.

The plan requires the vendor to inspect in accordance with product usage requirements.

The plan utilizes only existing paperwork and requires very little additional effort for its effective usage.

The plan is completely uniform with respect to all vendors, and serves to reduce judgement factors in rating vendors.

The plan provides objective quality evidence, gathered in category and quantity in accordance with the importance of the item. This information is of primary value in engineering development work.

PART NO. 5148734-1

VENDOR ABC MFG. CO.

Form 27-047 (10-54)

RECEIVING QUALITY CONTROL CHECK LIST AND HISTORICAL RECORD

NOT TO BE REMOVED FROM WORKING AREA

MERIT VALUE 28

VENDOR PART NO. (CROSS REFERENCE) None

VENDOR CODE 4932 CYCLE CONFORMITY XXX OF XXX

PART NAME Pressure Fitting Al., 356-T6

QUALIFIED PRODUCTS LIST NO. None

SPECIFICATIONS (PROCESS BULLETINS, ETC.)

TOOLS USED (SPECIAL AND STANDARD)

REF.:

PROCESS PER DRAWINGS, OPER. SHEET, AND P.O.

INSPECTION CHARACTERISTICS AND ACCEPTANCE QUALITY LEVEL

ITEM NO.	INSPECTION RECORD DATE	PURCHASE ORDER NO.	REC. REPORT NO.	LOT SIZE	1	2 HARDNESS TEST	3 MAGNETIC	4 VISUAL PER REMARK NOTE	5 PHYSICAL AND CHEMICAL REPORT	6 CYCLE CONFORMITY CHECK MADE	7 RELIEF PRESSURE	8 FLOW RATE	9 DIMENSIONAL B	23 DIMENSIONAL C	10	11	12	13	14	15	16	17	18	19	20 DASH NUMBER	21 INSPECTION STAMP	RECEIVING REJECTION REPORT NO.
						0.5% A	100% B			100% B	100% B	4% B	10% C													
1	1/29	099-123457	149899	100			5			9	6	20														364327
	2/10	099-123450	149900	200																						
	2/26	099-123455	149905	130							2															470593
	3/10	099-123460	149907	215						7		4	1													470930
5	4/1	099-123500	150010	100			2																			472000

FIGURE 1

NAI CERTIFICATION SHEET

VENDOR Alpha Mfg. Co. VENDEE NAI P.O.# _____

PART NAME Pressure Fitting PART # 51A8734-1 DATE 2/9/56

CHARACTERISTIC		Relief Pressure		Relief Flow					
TOLERANCE		580-600 psi		5-7 gpm					
SERIAL # \ CLASS		B		B					
1	(55)	583	(578)	6.2	6.3				
2	(56)	590	587	6.9	(7.1)				
3	(57)	587	586	7.0	(7.3)				
4	58	592	590	6.9	7.0				
5	59	590	592	6.1	6.3				
6	60	585	583	5.8	6.0				
7	61	592	590	6.3	6.4				
8	(62)	581	(575)	5.1	5.3				
9	(63)	581	(579)	5.8	5.6				
10	(64)	584	585	5.0	(4.9)				
11	65	590	592	6.2	6.3				
12	66	587	585	6.7	6.5				
13	(67)	596	599	7.0	(7.1)				
14	68	591	590	6.8	6.6				
15	(69)	580	(578)	5.1	5.0				
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
27									
28									
29									
30									
TOTAL									
AVERAGE									
RANGE									

FIGURE 2

NORTHROP AIRCRAFT, INC.
HAWTHORNE, CALIFORNIA
Receiving Report
FORM TS-24 (R. 7-56)

SHIP ALL MATERIAL
TO— 3901 S. MARIPOSA AVE., EL SEGUNDO, CALIF. PAGE OF
N.A.I. AT

P.O. NO. 099-1234567
DATE 2/9/56

VEND.

Alpha Mfg. Co. 4932
09876 El Segundo Blvd.
El Segundo, California
MERIT VALUE 28

CON. FIRM 1/6/56
☒ RESALE PERMIT AB-34874
☐ NOT RESALE
REQ. OR BID NO. 0-1234567-123
ENG. USE SHIP EFF. 5409-5412
PA. PR. CODE(S) 1234-0012345

ACCT. SALES ORDER 1234-123
CON. TRACT AF 99(000)-12345
TERMS
POE

SHIP VIA Air Ex.

ITEM	QUAN.	UNIT	N.A.I. PART NO., DESCRIPTION	UNIT PRICE	ITEM TOTAL
1	100	ea.	5148734-1 Fitting, pressure, Al. 356-T6 Make per "B" Chg. Dwg., No E.O.s effective		

INTERNAL DESTINATION DEPT. NO. ZONE ATT'N.

№ 272728

THE FOLLOWING ATTACHMENTS ARE INCORPORATED HEREIN AND MADE PART OF THIS ORDER BY REFERENCE

SCHED. B NO. S APPLY

I CERTIFY THE MATERIAL BELOW WAS RECEIVED, INSPECTED AND COMPLIES W/SPECIFICATIONS OR IS REJECTED AS SHOWN

ITEM	DATE	Q. REC'D.	Q. GOOD	R.R. No.	PACK. LIST	LOCATION	BY	REMARKS	TOT. REC'D.
1	2/9	100	60	5786	4321	xxx987			

BUYER - FLASH | AUDITED BY | DEPT. | AUDITED BY | DEPT. | BY INSP. | QTY. REJ. | REJ. REPT. NO.

APPROVED QUALITY CONTROL INSPECTION | APPLIES ONLY WHEN SIGNED | APPROVED U.S.A.F. CONTRACT OFFICER

FIGURE 3

RECEIVING REJECTION REPORT

Nº 364327

NORTHROP AIRCRAFT, INC.
HAWTHORNE, CALIFORNIA

MERIT VALUE 28

DATE 2/9/56

VENDOR	Alpha Mfg. Co.
ADDRESS	09876 El Segundo Blvd., El Segundo, California
PART NAME	Fitting, Pressure

REASON FOR REJECTION (EXPLICIT AND COMPLETE):

VIS B3 V	A-5 pcs are as described in MRDR #12345
DIM B2	B-20 pcs have undersize threads .007 PD
FUNCT B1	C-15 pcs not to spec. for Relief Pres. and Relief Flow

REJECTING INSPECTOR

VENDOR CODE NUMBER	4932
PART NO.	5148734-1
PURCHASE ORDER NO.	099-1234567
RECEIVING REPORT NO.	149899
ITEM NO.	1
CONTRACT/ACCOUNT NO.	1234-123
QUANTITY RECEIVED	100
QUANTITY REJECTED	40

SOURCE INSPECTED ☐
FIRST ARTICLE OR CHECK CASTING ☐
TOOLING ☐

TYPE OF INSPECTION
SAMPLING ☐ 100% ☒

REASON FOR REJECTION CODE

DISPOSITION

REWORK INSTRUCTIONS:

A- Parts OK as is per MRDR #12345

B- Not reworkable, scrap at NAI

C- Return to vendor for rework

QUANTITY	DISPOSITION	
35	REJECTED	☒
15	REWORKABLE	☒
5	ACCEPTED AS IS	☒
	REWORKABLE TO SPECIFICATION	☐

APPROVED	DATE
M. E. ENGINEER	
M. E. INSPECTOR	
CUSTOMER'S AGENT	

PROCUREMENT AGENCY ACTION

EXPLANATION:

CHARGE VENDOR	☒	CHARGE NORTHROP	☐
RETURN TO VENDOR	☒	SCRAP AT NORTHROP	☒
REPLACE ON ORIGINAL P. O.	☐	REWORK AT NORTHROP	☐
SUPPLEMENTARY P. O. ISSUED	☐	REPLACE AT NORTHROP	☐

VENDOR ERROR ☐
NORTHROP ERROR ☐
OTHER'S ERROR (INDICATE) ☐

PROCUREMENT AGENCY COPY

BY _____ PROCUREMENT AGENCY _____ DATE

FIGURE 4

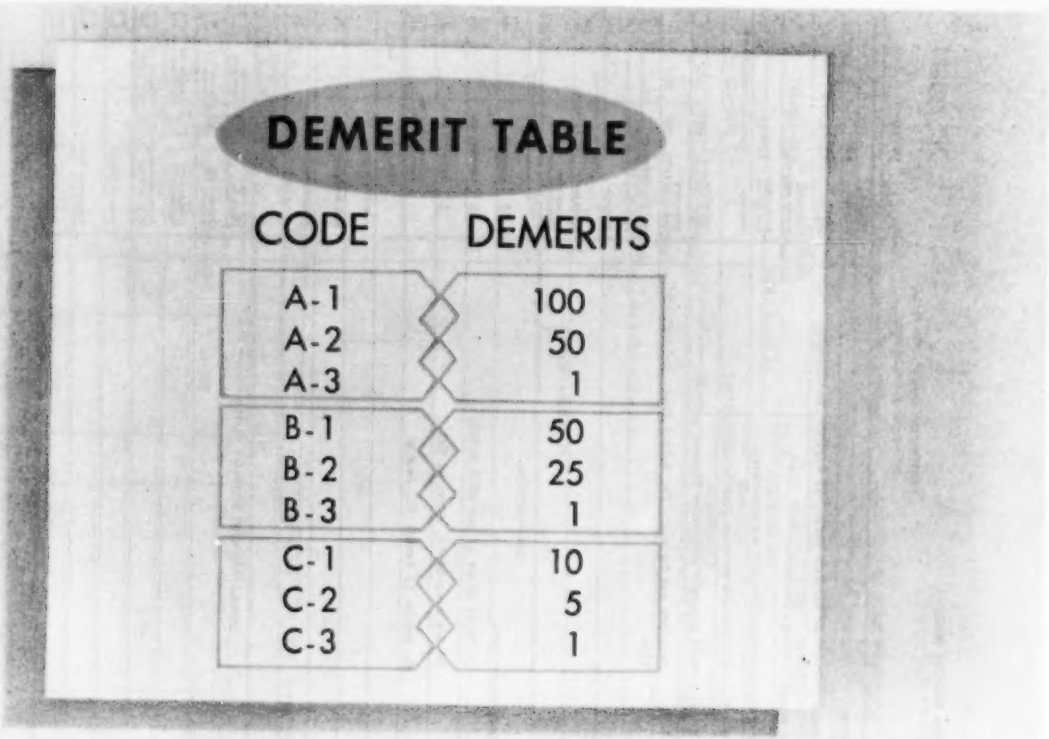
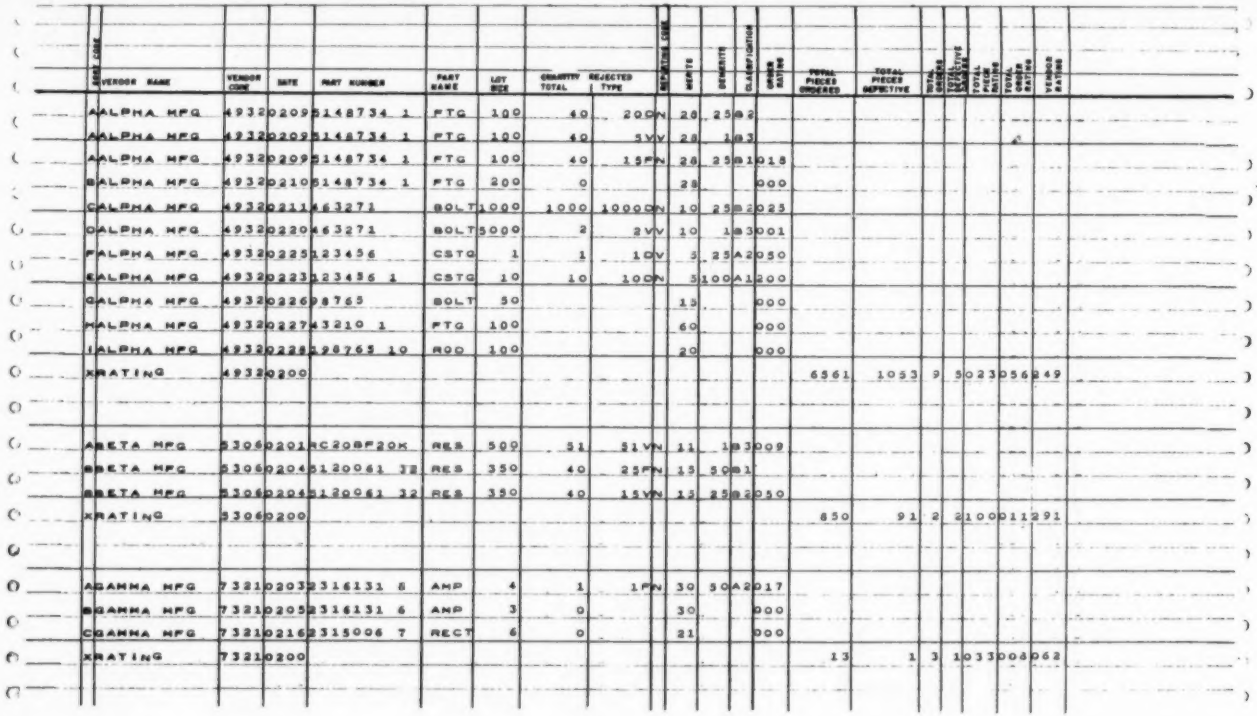
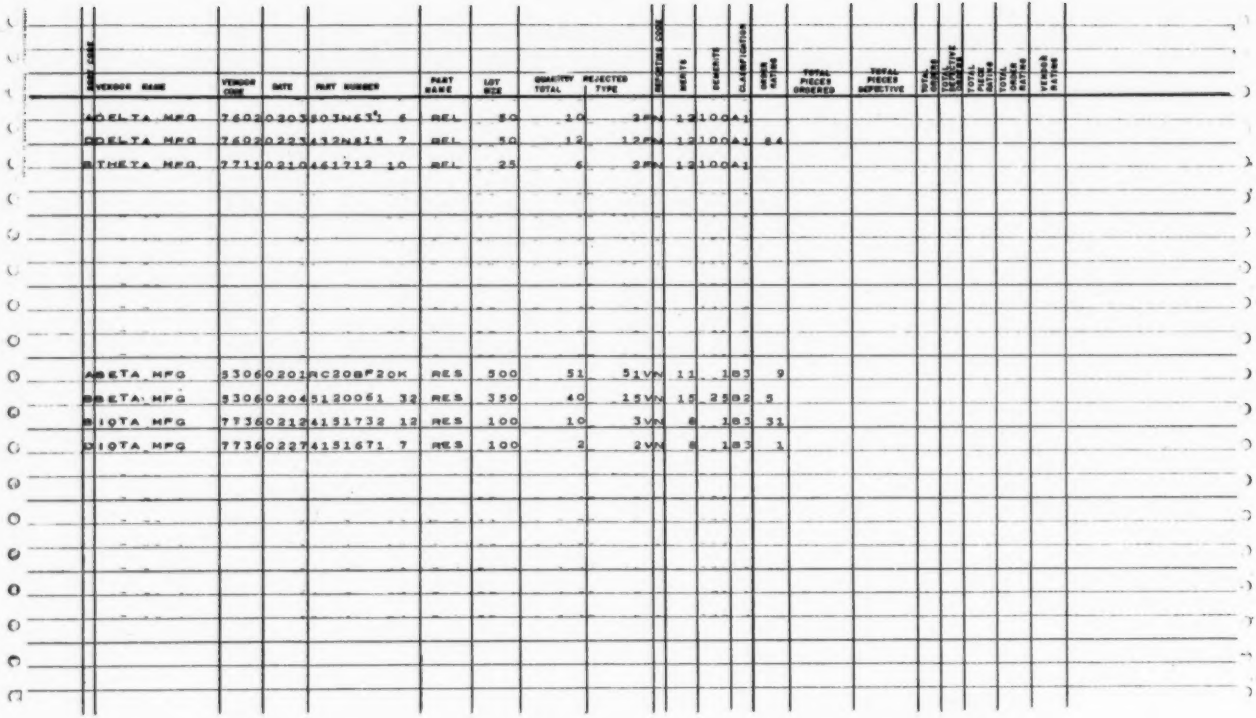

DEMERIT TABLE
CODE
DEMERITS
A-1
100
A-2
50
A-3
1
B-1
50
B-2
25
B-3
1
C-1
10
C-2
5
C-3
1

FIGURE 5



ASSY CODE	VENDOR NAME	VENDOR CODE	DATE	PART NUMBER	PART NAME	LOT SIZE	QUANTITY TOTAL	REJECTED TYPE	REJECTION CODE	WEEKS	DEMERITS	CLASSIFICATION	ORDER RATING	TOTAL PIECES ORDERED	TOTAL PIECES INSPECTIVE	TOTAL REJECTIVES	LOT REJECTIVE	TOTAL PIECE RATING	ORDER RATING	VENDOR RATING
A	ALPHA MFG	49320209	5148734 1	FTG	100	40	20	ON	28	25	B2									
A	ALPHA MFG	49320209	5148734 1	FTG	100	40	5	VV	28		1B3									
A	ALPHA MFG	49320209	5148734 1	FTG	100	40	15	FN	28	25	B1	018								
B	ALPHA MFG	49320210	5148734 1	FTG	200	0			28			000								
C	ALPHA MFG	49320211	463271	BOLT	1000	1000	1000	ON	10	25	B2	025								
D	ALPHA MFG	49320220	463271	BOLT	5000	2	2	VV	10		1B3	001								
F	ALPHA MFG	49320225	123456	CSTG	1	1	1	OV	5	25	A2	050								
E	ALPHA MFG	49320223	123456 1	CSTG	10	10	10	ON	5	100	A1	200								
G	ALPHA MFG	49320226	98765	BOLT	50				15			000								
H	ALPHA MFG	49320227	43210 1	FTG	100				60			000								
I	ALPHA MFG	49320228	198765 10	ROD	100				20			000								
X	RATING	49320200												6561	1053	9	50	23	056	249
A	BETA MFG	53060201	RC20BF20K	RES	500	51	51	VN	11		1B3	009								
B	BETA MFG	53060204	5120061 32	RES	350	40	25	FN	15	50	B1									
B	BETA MFG	53060204	5120061 32	RES	350	40	15	VN	15	25	B2	050								
X	RATING	53060200												850	91	2	2	100	011	291
A	GAMMA MFG	73210203	2316131 8	AMP	4	1	1	FN	30	50	A2	017								
B	GAMMA MFG	73210205	2316131 6	AMP	3	0			30			000								
C	GAMMA MFG	73210216	2315006 7	RECT	6	0			21			000								
X	RATING	73210200												13	1	3	1	033	008	062

FIGURE 6



RECORD CODE	VENDOR NAME	VENDOR CODE	DATE	PART NUMBER	PART NAME	LOT SIZE	QUANTITY TOTAL	REJECTED TYPE	INSPECTION CODE	MERITS	DEMERITS	CLASSIFICATION	ORDER RATING	TOTAL PIECES ORDERED	TOTAL PIECES DEFECTIVE	TOTAL ORDERS	TOTAL DEFECTIVE ORDERS	TOTAL PIECE RATING	TOTAL ORDER RATING	VENDOR RATING
A	DELTA MFG	7602	0203	503N63 6	REL	50	10	3	PN	12	100	A1								
D	DELTA MFG	7602	0223	432N815 7	REL	50	12	12	PN	12	100	A1	84							
B	THETA MFG	7711	0210	461712 10	REL	25	6	2	PN	12	100	A1								
A	BETA MFG	5306	0201	RC20BF20K	RES	500	51	51	VN	11	183		9							
B	BETA MFG	5306	0204	5120061 32	RES	350	40	15	VN	15	2582		5							
B	IOTA MFG	7736	0212	41S1732 12	RES	100	10	3	VN	8	183		31							
D	IOTA MFG	7736	0227	41S1671 7	RES	100	2	2	VN	8	183		1							

FIGURE 7

AQLs ARE HALF THE STORY

Edward R. Clark
Detroit Transmission Division
General Motors Corporation

The initial setting of an AQL for acceptance purposes must, of necessity, be arbitrary because of a lack of information on which to base such a decision. Indeed, the determination of a proper AQL is one of the knottiest problems faced in establishing a sampling plan. Many companies have established standard AQLs for various classifications of characteristics or parts. There have been some rather complicated methods derived for determining AQLs.

This paper demonstrates a method for evaluating the selected AQL regardless of how it was initially chosen. Several other points regarding the relationship of AQLs to incoming quality levels will be made.

The optimum sampling plan is one which maximizes protection of the customer at minimum inspection cost to the customer. As preliminary factors, these are extremely important, but any sampling plan must be reviewed periodically to determine its compatibility with the quality level of the material being submitted to it.

In the preceding discussion, references have been made to "AQLs". In most of the technical Quality Control literature, the term AQL stands for "Acceptable Quality Level". Mil-Std-105 defines AQL as "a nominal value expressed in terms of percent defective or defects per hundred units, whichever is applicable, specified for a given group of defects of a product". The Supply & Logisties Handbook H 105 indicates that the interpretation of AQL is sometimes made as the poorest tolerable average quality level of submitted product. Dr. Grant also establishes this latter concept of Acceptable Quality Level. In any case, it is evident that assignment of an AQL is the consumer's way of telling the vendor the average maximum amount of non-conformance he is willing to accept. In all of the literature concerning Acceptable Quality Levels, it is made equally clear that the vendor must do the best job he can because contractual obligations either imply or establish the vendors' responsibility for conformance to specification on all parts. However, once an Acceptable Quality Level has been established or assigned, it is rarely reviewed to determine whether or not it is realistic and whether or not the vendor is making a real attempt to improve the quality of submitted product.

In addition to the Acceptable Quality Level, there are two other AQLs that must be considered in any evaluation of a vendor's performance. These two additional AQLs are

The Actual Quality Level
and
The Attainable Quality Level

The Actual Quality Level is the vendor's performance over a series of lots generally in terms of percent defective. It is often referred to as the process average. The term Actual Quality Level has been substituted to keep the importance of its relationship to the other two AQLs clearly in mind.

The Attainable Quality Level is the level at which the vendor could

send in product if it were not for the presence of significantly inferior lots or inconsistent performance. It may be thought of as his process capability in percent defective.

The three types of AQLs then, that will be discussed in this example are

1. Acceptable Quality Level
   (What the customer wants)
   Symbol: Acc QL
2. Actual Quality Level
   (What the vendor is doing)
   Symbol: Act QL
3. Attainable Quality Level
   (What the vendor can do)
   Symbol: Att QL.

Tabulated below is a record of a series of 25 lots of the same part sampled and inspected. The Acceptable Quality Level is 4.0% and the sample size was 250. The reject number is 20.

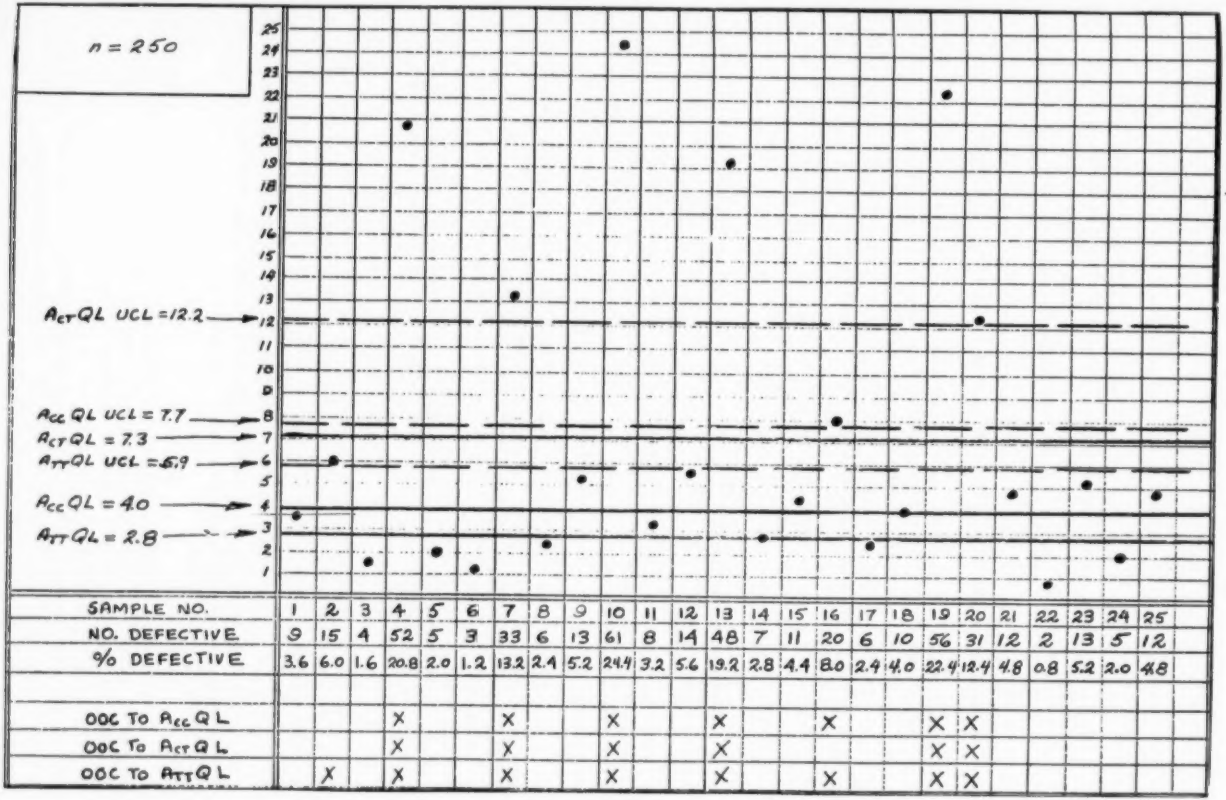
Lot No.	No. Defective	% Defective	Disposition
1	9	3.6	A
2	15	6.0	A
3	4	1.6	A
4	52	20.8	R
5	5	2.0	A
6	3	1.2	A
7	33	13.2	R
8	6	2.4	A
9	13	5.2	A
10	61	24.4	R
11	8	3.2	A
12	14	5.6	A
13	48	19.2	R
14	7	2.8	A
15	11	4.4	A
16	20	8.0	R
17	6	2.4	A
18	10	4.0	A
19	56	22.4	R
20	31	12.4	R
21	12	4.8	A
22	2	0.8	A
23	13	5.2	A
24	5	2.0	A
25	12	4.8	A

This data has been plotted on a p chart in Figure 1.

The Acceptable Quality Level of 4.0% is shown on this chart along with its 3-sigma Upper Control Limit.

The Actual Quality Level calculates to 456/6250 or 7.3% and is also shown on the chart with its 3-sigma Upper Control Limit.

The Attainable Quality Level, calculated in the usual manner by discarding points out of control to the Actual Quality Level, is 2.8% and



SAMPLE NO.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
NO. DEFECTIVE	9	15	4	52	5	3	33	6	13	61	8	14	48	7	11	20	6	10	56	31	12	2	13	5	12
% DEFECTIVE	3.6	6.0	1.6	20.8	2.0	1.2	13.2	2.4	5.2	24.4	3.2	5.6	19.2	2.8	4.4	8.0	2.4	4.0	22.4	12.4	4.8	0.8	5.2	2.0	4.8
OOC TO Acc QL				X			X			X			X			X			X	X				
OOC TO Act QL				X			X			X			X						X	X				
OOC TO Att QL		X		X			X			X			X			X			X	X				

its 3-sigma Upper Control Limit is also shown. Note that Lot 2 is out of control to the Attainable Quality Level of 2.8%. This suggests that the Attainable Quality Level may be even somewhat lower than 2.8% but further calculation demonstrates no significant lowering of the 2.8% with the available data.

This analysis may be summarized and tabulated as follows:

	%	No. of Lots O.O.C.	Q.L. Consistency
Acceptable Q.L.	4.0	7	72%
Actual Q.L.	7.3	6	76%
Attainable Q.L.	2.8	8	68%

This table of results also shows the number of lots out-of-control to each of the 3 Quality Levels and introduces, in the last column, the term Quality Level Consistency which is the % of lots consistent to each of the various Quality Levels. Note that although there is considerable difference in the 3 AQLs, the consistency levels are relatively stable. This indicates that not only is the quality performance of this vendor inconsistent to the Acceptable Quality Level, but it is equally inconsistent to his Actual Quality Level and to the inherent Attainable Quality Level of 2.8%.

Statistics without action have very little practical value for any of us. It is necessary that the data compiled above becomes the basis for an effective plan of correction. It is clear that the inconsistent performance of this vendor is causing the submission of lots considerably in excess of the desired quality level. It is also evident that the vendor's process is capable of producing at a quality level much superior to the originally specified Acceptable Quality Level.

The first course of action, therefore, is to change the Acceptable Quality Level from 4.0% to 2.8% or perhaps 3.0% if some standard sampling plans are being used. This alteration will best be represented by a reduction in rejection number from 20 to 15 rather than an increase in sample size so that the consumer is not affected in terms of inspection cost. Such a change in the sampling plan will result in an increased number of rejected lots unless action is taken with the vendor to improve the consistency of his process.

Analyses of the performances of many vendors (or departments in a plant) will demonstrate that inconsistency is the most serious quality problem existent. In other words, the poor operation of a process is a more prevalent cause of poor quality than the process itself. Most of us in Quality Control have made process capability studies and have found many times that even though a machine might not be capable of holding a required tolerance, it usually can be operated and controlled much better than it is.

In order to effect the corrective action necessary to make the vendor's submitted quality more consistent so that his Actual Quality Level will begin to approach his attainable Quality Level, a strict sequence of investigation is required to determine the assignable cause. This also occurs in sport. If a man is a 200 average bowler and he doesn't bowl 200 for several games, he is in a slump and we begin to look for what he is doing differently. There is no sudden jump to the conclusion that he is no longer a 200 bowler. In the case of the vendor,

it is apparent that he has not performed at a 2.8% level, but it is equally apparent that he can do it.

In order to determine the cause for inconsistent performance, the following sequence of action should be followed:

1. Determine, through a review of the records, which part characteristic defects caused the inconsistencies represented by those points above the Upper Control Limit for the Attainable Quality Level.

2. Review with the vendor each of these characteristics so that you and he are in agreement on what constitutes acceptability.

3. Gage or otherwise evaluate a sample of parts for each of these characteristics with the vendor so that he has a clear understanding of your interpretation of the characteristic (Step 2) and your physical evaluation of it.

4. Clearly identify the parts inspected in Step 3 and let the vendor take them to his plant so that he may determine whether or not his people can make the same decisions on each part. Defective parts, acceptable parts and borderline parts ought to be in the sample.

5. Check to determine that his final inspection is at least as tight or tighter than your's and that it includes very specific reference to the inconsistent defects.

6. Finally, the vendor should follow back to determine that the production operators have a clear knowledge of their responsibility through a review of Steps 3 and 4 with them.

Correction should be made at any step found unsatisfactory. The goal of such a program is to reduce the inconsistency of performance. Ideally, this goal will be realized when

$$\text{Acc QL} = \text{Act QL} = \text{Att QL}$$

If these 3 levels are equal, any further reduction will only be realized through process changes at the vendor's plant or engineering change at the customer's plant.

It is evident that this concept of Quality Level Consistency can form an excellent basis for Vendor Rating in conjunction with the various AQLs. It is evident that 2 vendors producing the same part may compare like this:

Vendor	AccQL	QLC	Act QL	QLC	Att QL	QLC
A	4.0	72	7.3	76	2.8	68
B	4.0	88	6.4	96	6.1	96

Essentially, it appears that Vendor A's problem is consistency of running his process while B's is the process itself.

The tightening of AQLs and improving the various acceptance plans adds to consumer quality protection but may seriously interfere with the consumer's quantity schedule. The acceptance function of any Receiving Inspection activity must be coupled with a planned program of correction and prevention of defects for optimum performance.

SAMPLING BY VARIABLES I: ONE-WAY PROTECTION ON $\bar{X}$

Max Astrachan
Air Force Institute of Technology

The principles of attributes inspection along with tables and procedures are well known and have been discussed in many places. Less well known, however, are techniques and tables for variables inspection.

It is the purpose of this series of four papers on sampling by variables, sponsored by the ASQC Committee on Education and Training, to discuss methods of constructing such plans, and the tables which are available in the literature. In the present paper, we consider the case of one-way protection on the mean, and show how to construct a variables plan in the case in which the product standard deviation is known.

In attributes inspection an item is classified either as defective or non-defective, or we count and record the number of defects in it. In variables inspection the characteristic in question is measured along a continuous scale in terms of inches, pounds, volts, seconds, etc. A decision relative to each piece inspected can be reached by comparing the measurement with the lower specification limit and/or the upper specification limit. This disposes of individual pieces. Since, however, we want to reach a decision about a lot on the basis of examining only a sample, this must be done by using information computed from the sample.

It is obvious that measurement of a quality characteristic gives much more information about an item than merely classifying it as defective or non-defective, or counting the number of defects in it. Hence, variables inspection of a sample yields more information about the quality of the lot than does attributes inspection. Practically, this results in reduced sample sizes required for specified degrees of protection, and is the major advantage of variables inspection. It should be considered as an alternative, therefore, when the cost of inspection by attributes is high.

Against this is the obvious disadvantage of the greater skill required to make the measurement, the cost of tools and gages, and the time required to make the necessary calculations. The latter is not as serious as would be supposed at first glance, however, since techniques have been devised for shortening them.

The use of variables plans requires a rather strong assumption about the nature of the distribution of the quality characteristic under consideration, viz., that it be normal. The frequency distribution of many measurements is roughly normal and hence from the practical point of view this assumption may be considered valid. However, it should be checked whenever possible. Throughout this paper we shall assume the normality condition satisfied.

As in many other situations, the choice between attributes and variables inspection, if there is one, must be made on economic grounds.

We shall illustrate the method of constructing a variables sampling plan by the following example:

Example 1 For a certain type of metal casting the lower specification limit (LSL) for tensile strength has been set at 55,000 psi. From previous experience it is known that tensile strengths of such castings are normally distributed with a standard deviation of 3500 psi, i.e., $\sigma' = 3500$. The variables plan consists of the determination of a sample size n, and a constant K, and operates as follows:

a. Select a random sample of size n from the lot.

b. Find $\bar{X}$, the mean tensile strength of the sample.

c. If $\bar{X}$ is greater than or equal to K, accept the lot; if not, reject it.

Clearly, if the lot mean $\bar{X}'$ is large enough there will be very few castings with tensile strengths below 55,000 psi. Suppose $\bar{X}' = 65,000$. Then the percent below the specification limit can be determined by first finding the normal deviate

$$\frac{LSL - \bar{X}'}{\sigma'} = \frac{55,000 - 65,000}{3500} = -2.86,$$

and then referring to a table of areas under the normal curve. We find that 0.21% will not meet the specification.

If the lot mean is not much greater than 55,000 psi, the percent of defective castings will be large. Suppose $\bar{X}' = 61,000$. As above, then,

$$\frac{LSL - \bar{X}'}{\sigma'} = \frac{55,000 - 61,000}{3500} = -1.71,$$

and from a table of areas under the normal curve, 4.36% of the castings will be below the specification limit.

Suppose then that we consider 65,000 psi as an acceptable quality level for $\bar{X}'$ and 61,000 psi as a rejectable quality level. With each we associate risks of rejection and acceptance, say $\alpha = .02$ and $\beta = .06$, respectively. This means then that the probability of rejecting lots of acceptable mean quality ($\bar{X}' = 65,000$) is .02, and the probability of accepting lots of rejectable mean quality ($\bar{X}' = 61,000$) is .06. As in all cases, whether or not these quantities (levels and risks) are satisfactory must be decided by management. Since our test is a test on means, we shall be working with the distribution of means, which we know is normal. Figure 1 shows the distribution of $\bar{X}$ for $\bar{X}' = 61,000$ and $\bar{X}' = 65,000$ together with the corresponding α and β.

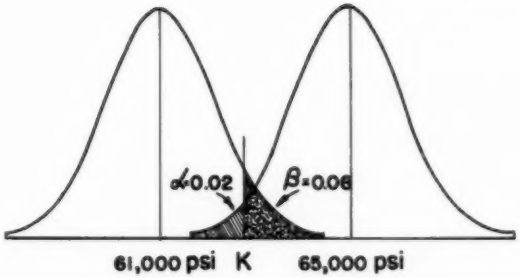


Fig. 1. Distribution of $\bar{X}$ for acceptable and rejectable quality levels, and the corresponding risks α and β.

To determine the values of n and K, we proceed as follows: Since we have agreed to reject acceptable lots (with $\bar{X}' = 65{,}000$) 2% of the time, we may write*

$$\frac{K - 65{,}000}{\frac{3500}{\sqrt{n}}} = -2.054 \qquad (1)$$

Further, to accept lots of rejectable quality (with $\bar{X}' = 61{,}000$) 6% of the time implies that*

$$\frac{K - 61{,}000}{\frac{3500}{\sqrt{n}}} = +1.555 \qquad (2)$$

The constants on the right in (1) and (2) are determined from the standardized normal curve so as to give $\alpha = .02$ and $\beta = .06$.

* Remember that we are working here with the distribution of the mean, for which the standard deviation is $\sigma/\sqrt{n}$.

Subtracting (1) from (2) gives

$$\frac{4000}{\frac{3500}{\sqrt{n}}} = 3.609,$$

whence

$$4000 = 3.609\left(\frac{3500}{\sqrt{n}}\right)$$

$$\sqrt{n} = (3.609)(.875) = 3.16$$

$$n = 9.99.$$

So we would use a sample of size n = 10. (We always round our computed value up to the next integral value of n instead of to the nearest whole number.)

To determine K, we may substitute n = 10 in either (1) or (2). Using (1) would leave α at .02 but change β slightly, whereas substituting in (2) would leave β at .06 but change α slightly. Which we would do in a particular case depends upon management.

Using (1) we have

$$K = 65{,}000 - 2.054\left(\frac{3500}{\sqrt{n}}\right)$$

$$= 65{,}000 - 2.054(1106.8)$$

$$= 65{,}000 - 2{,}273$$

$$= 62{,}727$$

To see what happens to β, put this value of K in (2) to get

$$\frac{62{,}727 - 61{,}000}{\frac{3500}{\sqrt{10}}} = \frac{1727}{1106.8} = 1.560$$

as the normal deviate,

whence $\beta = .0594$ which is close to .06 because our computed value of n was very close to 10. Had it come out closer to 9 the value of β would have differed from .06 by a larger amount.

Had we used equation (2), the value of K would be 62,721, β would remain at .06, and α turns out to be .0198.

To summarize then, our plan operates as follows:

a. Select a random sample of 10 castings.

b. Find $\bar{X}$, the mean tensile strength of the sample.

c. If $\bar{X}$ is greater than or equal to 62,727, accept the lot, otherwise reject it.

If the lot mean is as great as 65,000 psi, the probability of accepting it is 0.98, and such lots will contain 0.21% castings with tensile strengths less than the lower specification limit of 55,000 psi. If the lot mean is as low as 61,000 psi, the probability of accepting the lot is 0.0594, and such lots will be 4.36% defective.

To see how the plan operates on lots of other mean values, we construct its operating characteristic curve. The essential calculations are shown in Table 1. Column (2) gives the normal deviate values for K relative to lot mean values in column (1) for the $\bar{X}$ distribution. The entries in column (3) are obtained from a table of areas under the normal curve and are the probabilities of accepting lots with mean values listed in column (1). The entries in column (4) are the normal deviate values for the lower specification limit in the distribution of individuals, relative to the lot mean values in column (1). In column (5) are the proportions of defective castings (i.e. with tensile strengths less than 55,000 psi) in the lots of stated mean values in column (1).

Table 1 Calculations for the OC curve for the sampling plan in Example 1. ($n = 10$, $K = 62,727$)

$\bar{X}'$	$z = \frac{K - \bar{X}'}{3500/\sqrt{n}}$	P_a = area above z in (2)	$z = \frac{55,000 - \bar{X}'}{3500}$	Area below z in (4)
(1)	(2)	(3)	(4)	(5)
59,000	3.37	.0004	-1.14	.1271
60,000	2.46	.0069	-1.43	.0764
61,000	1.56	.0594	-1.71	.0436
62,000	.66	.2546	-2.00	.0228
63,000	- .25	.5987	-2.29	.0110
64,000	-1.15	.8749	-2.57	.0051
65,000	-2.05	.9798	-2.86	.0021
66,000	-2.96	.9985	-3.14	.0008
67,000	-3.86	1.0000	-3.43	.0003

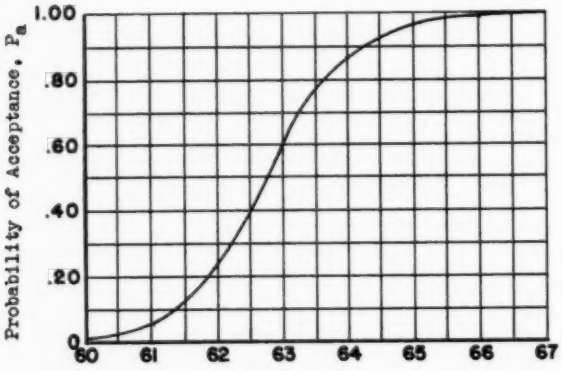


Fig. 2. $\bar{X}'$, tensile strength in thousands of psi.

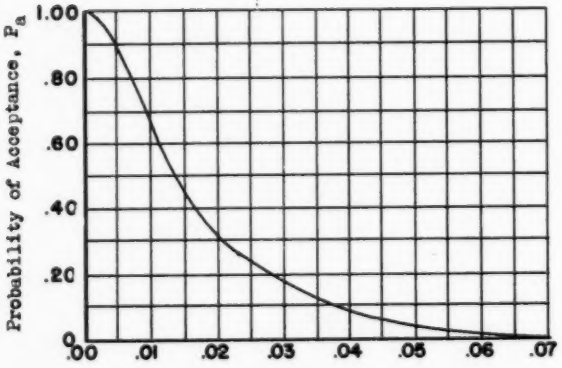


Fig. 3. Proportion of castings below 55,000 psi.

Figure 2 shows the operating characteristic curve for the plan, obtained by plotting column (3) against column (1). We can determine from the curve the probability of accepting lots of various mean values, if they are offered. In Figure 3 we have plotted column (3) against column (5). This curve shows the probability of accepting lots of various incoming fraction defectives (fraction of pieces below 55,000 psi).

To show how a variables plan can be constructed for the case in which an upper specification limit (USL) is given, consider

Example 2 For a certain type of metal casting the upper specification limit has been set at 90,000 psi. From previous experience it is known that tensile strengths of such castings are normally distributed with $\sigma' = 3500$ psi. We want to find the sample size n and constant K so that our procedure will be as follows:

a. Select a random sample of size n from the lot.

b. Find $\bar{X}$, the mean tensile strength of the sample.

c. If $\bar{X}$ is less than or equal to K, accept the lot, otherwise reject it.

Here we want to guard against the lot mean being too high rather than too low, as was the case in the first example. Let us consider 80,000 psi as an acceptable value for the lot mean, and 84,000 psi as unacceptable, with risks of rejection as $\alpha = .02$, and of acceptance as $\beta = .06$, respectively. Then proceeding as in Example 1, if $\bar{X}' = 80{,}000$, 0.21% of the castings will be above the upper specification limit, whereas if $\bar{X}' = 84{,}000$, 4.36% will be unacceptable.*

To find the values of n and K, we set up the two equations analogous to (1) and (2):

$$\frac{K - 80{,}000}{\frac{3500}{\sqrt{n}}} = +2.054 \qquad (3)$$

and

$$\frac{K - 84{,}000}{\frac{3500}{\sqrt{n}}} = -1.555. \qquad (4)$$

* These percentages turn out to be the same as in Example 1 only because of our selection of $\bar{X}' = 80{,}000$ and $\bar{X}' = 84{,}000$ as the acceptable and rejectable values. We could, of course, have used other suitable values. Using $\alpha = .02$ and $\beta = .06$ as before, also duplicates some of the following numerical work.

Solving for n, gives as before,

$$4000 = 3.609\left(\frac{3500}{\sqrt{n}}\right)$$

whence

$$n = 9.99$$

and we shall use a sample size of 10. To find K we substitute this value in (3) to leave α unchanged. This gives

$$K = 80{,}000 + 2.054\left(\frac{3500}{\sqrt{10}}\right)$$

$$= 80{,}000 + 2{,}273$$

$$= 82{,}273.$$

The value of β, however, is reduced to .0594 as in Example 1. Had we used (4) to find K, its value would be 82,279, β would remain at the specified value of .06, but α would be reduced to .0198.

Our procedure to determine the disposition of a lot then, is to first find the mean breaking strength of a random sample of 10 castings. If the sample mean is less than or equal to 82,273 psi, accept the lot, otherwise reject it.

The attributes plan corresponding to the variables plans derived above would require a large sample. Thus, for the situation in Example 1, each casting would have to be subjected to a 55,000 psi tensile test and rated as to whether or not it breaks. The use of the much smaller sample size in the variables test results in considerable savings in time and money.

The use of the plan in Example 1 can be illustrated by throwing four dice and noting the total number of dots. Since for four dice $\bar{X}' = 14.0$ and $\sigma' = 3.42$, if we call each dot 1000 psi, we can simulate a population for which $\sigma' = 3420$, which is close to our value of 3500 psi. Calling the total number of dots observed on a throw as the number of thousands of psi above 50,000, for example, we shall be sampling from a population for which $\bar{X}' = 64{,}000$. Rolling the dice n = 10 times should produce a mean which will be greater than or equal to our K value of 62,727, and thus lead to acceptance, about 87.5% of the time as can be seen from Table 1 or Figure 2. Similarly, we can adjust $\bar{X}'$ to other values and try the plan.

One other final comment might be made. Because of the relationship between the mean of a normal distribution and the fraction defective, it is possible to construct variables plans instead of attributes plans to control the latter. It is only necessary to convert from fractions

defective to corresponding values of $\bar{X}'$. Thus Example 1 could have been restated as follows: The lower specification limit on tensile strength is 55,000 psi. Design a variables sampling plan which will have an AQL = 0.0021, lot tolerance fraction defective $p_t = 0.0436$, $\alpha = 0.02$, and $\beta = 0.06$.

SAMPLING BY VARIABLES II: TWO WAY PROTECTION ON $\bar{X}'$

Fred C. Leone
Case Institute of Technology

In the preceding discussion we considered a variables sampling plan where the consumer was interested in a one way protection. We illustrated the exact nature of the plan by an example showing a step by step procedure for its construction. Finally, we followed this with a detailed consideration of the operating characteristic curve.

Suppose we extend this idea to two sided protection. It is not difficult to visualize a number of cases where a product can have too high as well as too low a value to satisfy the consumer's specifications. Let us consider the problem presented in the earlier discussion.

For a metal casting both a lower and upper specification limit have been prescribed. Suppose these are 55,000 psi and 90,000 psi respectively. Again assume that the tensile strengths are normally distributed with $\sigma' = 3500$ psi. For the lower specification a lot with a mean of 65,000 psi is considered to be acceptable quality. One with a mean of 61,000 psi is of rejectable quality. On the upper specification this acceptable and rejectable quality is 80,000 psi and 84,000 psi respectively. As in the example presented in the earlier discussion, let us agree to have a risk of rejecting acceptable quality equal to .02 , and of accepting rejectable quality equal to .06. That is $\alpha = .02$ and $\beta = .06$. Let this .02 and .06 be equally divided at both specification limits. This means $\alpha/2 = .01$ and $\beta/2 = .03$ at each end. Graphically this can be represented as Figure 1.

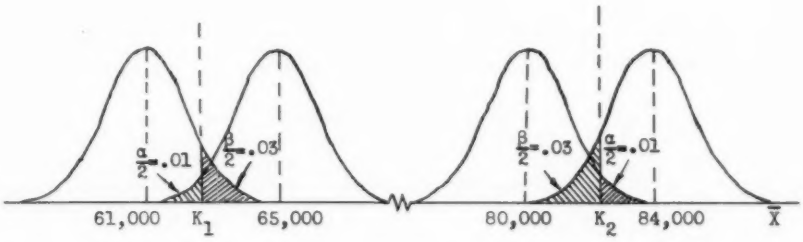


Fig. 1. Distribution of $\bar{X}$ for Acceptable and Rejectable Quality Levels.

Remember that the values of the $\bar{X}'$'s were chosen so that a particular percentage of the product was outside of the specification limits. To determine the values of K_1, K_2 and n the following equations are used:

$$(1) \qquad \frac{K_1 - 65{,}000}{3500/\sqrt{n}} = -2.326$$

(2) $$\frac{K_1 - 61,000}{3500/\sqrt{n}} = 1.881$$

(3) $$\frac{K_2 - 84,000}{3500/\sqrt{n}} = -2.326$$

The constants on the right of equations (1), (2) and (3) are determined so as to give $\alpha/2 = .01$ and $\beta/2 = .03$ from the standardized normal curve. A fourth equation which can be derived from the first three, and therefore is not independent* of these, is

(4) $$\frac{K_2 - 80,000}{3500/\sqrt{n}} = 1.881$$

Solving for n from equations (1) and (2) we obtain

$$4000 = \frac{4.207 \times 3500}{\sqrt{n}}$$

$$\sqrt{n} = \frac{4.207 \times 7}{8}$$

$$n = 13.6$$

$$n = 14 \quad \text{(always rounded up)}$$

If we substitute $n = 14$ in equation (1) we get

$$K_1 = \frac{3500\ (-2.326)}{\sqrt{14}} + 65,000$$

$$= 62,824$$

From equation (3)

$$K_2 = \frac{3500\ (-2.326)}{\sqrt{14}} + 84,000$$

$$= 81,824$$

This gives $\alpha = .02$. To determine the value of β, substitute the value of K_1 and n in the left of equation (2). The result gives us a value of z for the standardized normal distribution. The area to the right of z is then the calculated $\beta/2$. We do not expect this to give us

*Equation (4) is obtained by subtracting equation (1) from the sum of equations (2) and (3).

exactly $\beta/2 = .05$ since the value of n was rounded to a whole number. To continue

$$\frac{K_1 - 61{,}000}{3500/\sqrt{n}} = \frac{62{,}823 - 61{,}000}{3500/\sqrt{14}} = 1.949$$

Therefore, from the normal curve

$$\beta/2 = .0256$$
$$\text{and } \beta = .051$$

If instead of substituting n = 14 in equations (1) and (3) we had used (2) and (4) we would obtain

$$K_1 = 62{,}760$$

$$K_2 = 81{,}760$$

$$\text{with } \alpha = .017$$

$$\beta = .06$$

We now have a choice of two sets of K_1 and K_2, the first giving $\alpha = .02$ and $\beta = .051$, and the second with $\alpha = .017$ and $\beta = .06$. The choice of one of these two is an administrative decision and need not concern us here. The former is more conservative, however. With the former plan the procedure is as follows:

(1) Select a random sample of size 14
(2) Find the mean tensile strength of this sample, namely, $\bar{X}$.
(3) If $\bar{X}$ is between 62,824 psi and 81,824 psi accept the lot. Otherwise, reject the lot.

This plan, then states that if a lot mean is 65,000 psi or 80,000 psi the probability of its being accepted is .99 (that is, $1-\alpha/2$) and such lots will have .21 percent of the castings with tensile strength below 55,000 psi in the former case and the same percent above a tensile strength of 90,000 psi in the latter case. If a lot mean is as low as 61,000 psi or as high as 84,000 psi the probability of its being accepted is .0255 (that is, $\beta/2$) and such lots will have 4.27 percent defective.

Consider now a modified example. A cylindrical part is produced for an assembly. The outer diameter is required to have an average quality of .4275 inches. If these parts are sampled we must be willing to have a risk of rejecting good quality as well as a risk of accepting bad quality. Suppose we decide that lots with an average quality below .4270 in. or above .4280 in. are not desirable. However, we are willing to accept those lots of small pieces, say 10 percent of the time, and those of large pieces, say 10 percent of the time also. This is our β, or the consumer's risk. On the other hand, we will agree that, say 5 percent of the time we will reject lots of good quality. This is the α, or producer's risk. Again, returning to the assumption of normality of the distribution of sample means, our problem may be displayed graphically as follows:

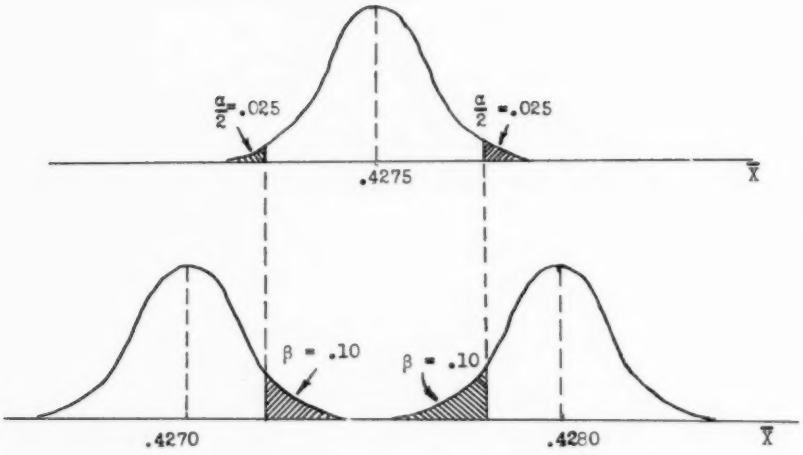


Fig. 2. Distribution of $\bar{X}$ for Acceptable and Rejectable Quality Levels.

Assume that σ' is known and is equal to .0003 in. Our problem then is as follows: Given an acceptable quality level of .4275, the lower and upper undesirable qualities of .4270 and .4280, and the standard deviation $\sigma' = .0003$, how large a sample should we take and what are the lower and upper acceptance limits K_1 and K_2 for the average of this sample of n pieces? We are willing to accept the product with a mean of .4270 ten percent of the time, and with a mean of .4280 ten percent of the time, and to reject the product at average quality level equal to .4275 five percent of the time. K_1, K_2 and n are the unknowns. The equations for the solution are:

$$\frac{K_1 - .4270}{.0003/\sqrt{n}} = 1.282 \tag{5}$$

$$\frac{K_2 - .4280}{.0003/\sqrt{n}} = -1.282 \tag{6}$$

$$\frac{K_1 - .4275}{.0003/\sqrt{n}} = -1.960 \tag{7}$$

The fourth equation, which can be derived from the first three is:

$$(8) \qquad \frac{K_2 - .4275}{.0003/\sqrt{n}} = 1.960$$

Solving for n from equations (5) and (7) we obtain:

$$.0005 = 3.242 \ \frac{.0003}{\sqrt{n}}$$

$$\sqrt{n} = 3.242 \times .6 = 2.00$$

$$n = 4$$

Substituting n = 4 in equation (5) and then in (6) yields

$$K_1 = .4272$$

$$K_2 = .4278$$

As a check, determine the α for an acceptable quality level of .4275 and acceptance lower and upper limits on the sample mean of .4272 and .4278 respectively. This turns out to be .05. The reason α is almost exactly the value originally required is that there was essentially no rounding off of the exact value of n calculated to obtain an integer.

In order to see the effectiveness of the sampling plan in accepting good lots and rejecting poor ones, we construct its operating characteristic curve. Table 1 gives the necessary calculations for points on the curve. Note that column (1) gives the various values of the lot mean, $\bar{X}'$. The values of z_1 and z_2 in the columns (2) and (3) respectively give the normal deviates for values of $\bar{X}'$. Column (4) gives the probability of acceptance of a lot whose mean is specified in column (1). These probabilities are obtained from a table of areas under the normal curve.

The operating characteristic curve of the above plan is given in Figure 3. In this the abscissa is the true mean $\bar{X}'$ in inches obtained from column (1) of Table 1. The ordinate is the probability of acceptance of a lot whose mean is $\bar{X}'$.

The above procedure could be employed for a sampling plan where the AQL, p_t', α and β are specified and the tolerance limits are given. Again, σ' is assumed known and the distribution is assumed normal. For each of the tolerance limits, the lower and upper tolerance means are computed and we proceed as before.

Table 1. Calculation for the OC Curve for the Sampling Plan in Example 2 ($K_1 = .4272$, $K_2 = .4278$, $n = 4$)

$\bar{X}'$	$z_1 = \frac{K_1 - \bar{X}'}{.0003/\sqrt{n}}$	$z_2 = \frac{K_2 - \bar{X}'}{.0003/\sqrt{n}}$	$P_{(acc)}$ = areas between z_1 and z_2
(1)	(2)	(3)	(4)
.4269	2.00	6.00	.0228
.4270	1.33	5.33	.0918
.4271	0.67	4.67	.2514
.4272	0.00	4.00	.5000
.4273	-0.67	3.33	.7482
.4274	-1.33	2.67	.9044
.4275	-2.00	2.00	.9545
.4276	-2.67	1.33	.9044
.4277	-3.33	0.67	.7482
.4278	-4.00	0.00	.5000
.4279	-4.67	-0.67	.2514
.4280	-5.33	-1.33	.0918
.4281	-6.00	-2.00	.0228

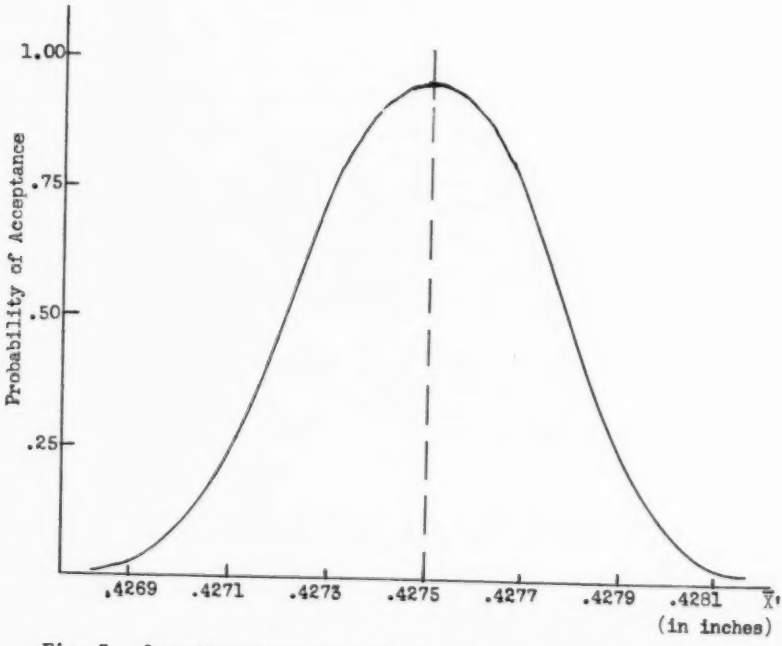


Fig. 3. Operating Characteristic Curve for the Double Limit Sampling Plan $n = 4$, $K_1 = .4272$ and $K_2 = .4278$.

It should again be stressed that a variables sampling plan requires smaller sample sizes than an equivalent attribute sampling plan. This is evident from the fact that in variables sampling, more than just "good" or "bad" is assigned to each item tested. At times one may not have a choice between attribute or variables sampling. But if a measurement can be taken on a quality characteristic of an item, and especially if the cost of testing is expensive, then the variables sampling plan will, in general, be more advantageous.

SAMPLING BY VARIABLES III - PROTECTION ON VARIABILITY

Gayle W. McElrath
University of Minnesota

Jacob E. Bearman
University of Minnesota

INTRODUCTION:

Basically, the disposition of a lot which is submitted to a variables acceptance procedure depends on one of two common quality criteria. When the quality being tested is measured, the question may be (1) is the lot different from a standard level or (2) has it excessive variability. Most often the acceptance criterion is based on the level of lot quality. However, occasionally, the criterion of process variability is the necessary quality characteristic for acceptance. This paper concerns itself with the latter.

A few examples will help realize the problem.

1. Machine tools might not operate successfully if the shipment of bar stock varies too greatly.

2. Time standards require a certain measure of consistency of performance rating among different engineers.

3. A problem of balance might require that a set of n objects do not have too great a variability in weight.

4. Matching painted panels requires that the characteristics of the surface finishing do not vary too greatly.

5. The variability in weight of specific units to be packaged must adhere to given specifications around a specified average.

6. The variability of the life of n tubes in an electronic device must not exceed a given value in order to obtain economic maintenance and replacement.

7. A set of operations must take place within a given time interval. Thus it is desired that the range of the longest and shortest interval of time not exceed a certain value.

STATEMENT OF THE PROBLEM:

It is heartening to learn that we can again use the concepts of acceptable quality level (AQL) and rejectable quality level (RQL) together with their respective producer's risk (P.R. or α) and consumer's risk (C.R. or β). That is, the same basic criteria which are ingredient in making decisions of attributes acceptance are still the important ingredients in determining the variables acceptance when the lot quality is measured by the variation. This measure of variation may be either the standard deviation or the range.

To be specific, consider that the acceptable quality level for variability is given by $\sigma'_{AQL} = 5$ measured in whatever units are appropriate. The producer's risk is some small quantity, say $\alpha = .01$ or 1%. That is, if lots of quality σ'_{AQL} or better are submitted, we will reject them on the average, at most, 1 time in 100. On the other hand, consider a lot is rejectable if the variability is $\sigma'_{RQL} = 10$. The consumer's risk is, say $\beta = .05$ or 5%. That is, if the quality of submitted lots is $\sigma'_{RQL} = 10$ or worse, we wish to accept these lots, at most, on the average of 1 time in 20. In addition, let us assume that the quality measure can be adequately represented for practical purposes by the normal frequency distribution.

Our objective, then, is to generate a decision making rule which gives the above desired protection. We shall find two numbers n and σ_a, where n is the sample size and σ_a is the acceptance standard deviation. That is, σ_a is the maximum value that the sample standard deviation σ can have and result in a decision to accept the lot. Briefly, our decision rule can be stated:

1. Take n observations.
2. If $\sigma \leq \sigma_a$, accept the lot, and if $\sigma > \sigma_a$, reject the lot

   where $$\sigma = \sqrt{\frac{\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}}{n}}$$

   and x_i, $i = 1,2,\ldots, n$ are the individual observations; and Σ is understood as the summation over all observations.

FIRST METHOD OF SOLUTION: Single Sampling Plan Based on Sample Standard Deviation σ.

Now we are ready to find both n and σ_a. Consider the following "cook book" directions:

1. Choose σ'_{AQL}, σ'_{RQL}, α, and β. (In the above example, $\sigma'_{AQL} = 5$, $\sigma'_{RQL} = 10$, $\alpha = .01$, $\beta = .05$).
2. Form the ratio $\left(\sigma'_{RQL}/\sigma'_{AQL}\right)^2 = \left(\frac{10}{5}\right)^2 = 4$.

   Use Table A for $\alpha = .01$.
3. Proceed down column (6) until you bracket the value 4 by two numbers, namely, 4.02 and 3.85.

4. From column (1), select the sample size $n = 18$ which corresponds to the smaller of the bracketing values.

5. From column (2) obtain the corresponding $\chi^2_a = 33.41$. Form the equation

$$\frac{n\sigma_a^2}{\sigma'^2_{AQL}} = \chi^2_a$$

and solve for σ_a. That is

$$\frac{18\sigma_a^2}{25} = 33.41$$

$$\sigma_a = \sqrt{\frac{33.41 \times 25}{18}} = 6.81$$

To summarize our decision rule:

1. Take $n = 18$ observations.
2. Calculate the sample standard deviation, σ.
3. If $\sigma \leq 6.81$ accept the lot, and
   if $\sigma > 6.81$ reject the lot.

Companion to every sampling plan should be an operating characteristic curve, usually called OC curve. We wish to determine the quantity $P(\sigma \leq \sigma_a | \sigma') = P_a$ (the probability of accepting a lot, given that σ' is the true standard deviation). Now $P(\sigma \leq \sigma_a | \sigma') = P(\sigma^2 \leq \sigma_a^2 | \sigma') = P\left(\frac{n\sigma^2}{\sigma'^2} \leq \frac{n\sigma_a^2}{\sigma'^2} \mid \sigma'\right) = P_a$. But for a given σ' the quantity $\frac{n\sigma^2}{\sigma'^2}$ has a χ^2 distributed with n-1 degrees of freedom.

Therefore, our basic problem is merely to substitute various values of σ' in the expression

$$P\left(\chi^2 \leq \frac{n\sigma_a^2}{\sigma'^2} \mid \sigma'\right) = P_a \quad \text{or}$$

$$1 - P\left(\chi^2 > \frac{n\sigma_a^2}{\sigma'^2} \mid \sigma'\right) = P_a.$$

The quantity P_a is determined from the χ^2 distribution. A very

useful χ^2 table for this purpose is available as Table C in "Engineering Statistics and Quality Control" by Irving Burr published by the McGraw-Hill Book Co., Inc., 1953.

TABLE I: Values for OC Curve
Single Sampling Plan for Variability
$\sigma'_{AQL} = 5$, $\sigma'_{RQL} = 10$, $\alpha = .01$, $\beta = .05$
Plan: $n = 18$; $\sigma \leq 6.81$ accept, $\sigma > 6.81$ reject.

σ'	$\frac{n\sigma_a^2}{\sigma'^2}$	P_a
0	∞	100
4	52.20	100
5	33.41	99
6	23.20	86
7	17.04	55
8	13.05	27
9	10.31	11
10	8.35	4
11	6.90	1.5
12	5.80	0.6

SECOND METHOD OF SOLUTION: Single Sampling Plan Based on Sample Sum of Squares SS.

It should be clear that the sum of squares, $SS = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$ is satisfactory as a statistic to determine acceptance for variability. Actually, to completely calculate the sample standard deviation does not add information. In fact, the only addition is unnecessary work which we may choose to omit. Steps (1), (2), and (3) of the "cook book" directions are the same as before. Step (4) is a bit different and easier. It is

4. Obtain χ^2_α from column (2); form the equation

$$n\sigma_a^2 = \chi^2_\alpha \cdot \sigma^2_{AQL}$$
$$= 33.41 \times 25 = 835.25$$

Now our decision rule becomes:

1. Take $n = 18$ observations.

2. Calculate the sample sum of squares

$$SS = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$$

where x_i, $i = 1,2,\ldots,$ n are the individual observations.

3. If $SS \leq n\sigma_a^2 = 835.25$, accept the lot, and
if $SS > 835.25$, reject the lot.

The OC curve for the plan based on the sample sum of squares SS is identical to that of the previous OC curve based on the sample standard deviation σ.

It is possible to choose the values of σ'_{AQL}, σ'_{RQL}, α, β such that the ratio $\left(\sigma'_{RQL}/\sigma'_{AQL}\right)^2$ is not included in Table A. That is, the sample size n required to give the specified protection is greater than the sample sizes listed. For an illustration of this situation see "Techniques of Statistical Analysis" by Eisenhart, Hastay and Wallis, McGraw-Hill Book Co., Inc., 1946, pp. 231-232.

THIRD METHOD OF SOLUTION: Single Sampling Plan Based on the Sample Range R.

At times it is much more convenient and practical to use the sample range R as the measure of sample variability. For small samples, say $n = 2$ to $n = 12$, the efficiency of the range is quite impressive.

Again it is our objective to select the quality levels of variability which define the plan. That is, suppose we choose $\sigma'_{AQL} = 5$, $\sigma'_{RQL} = 15$, $\alpha = .01$, and $\beta = .05$.

Now we must determine two numbers, the sample size n and the acceptance range R_a. That is, R_a is the maximum value that the sample range R can have and result in a decision to accept the lot. The decision rule becomes

1. Take n observations.
2. Determine the sample range R.
3. If $R \leq R_a$, accept the lot, and if $R > R_a$, reject the lot.

These are the "cook book" steps.

1. Determine σ'_{AQL}, σ'_{RQL}, α, and β.
(In our example $\sigma'_{AQL} = 5$, $\sigma'_{RQL} = 15$, $\alpha = .01$, and $\beta = .05$.)
2. Form the ratio, $\sigma'_{RQL}/\sigma'_{AQL} = (15/5) = 3$.
3. Use Table B for $\alpha = .01$. Proceed down column (6) until you bracket the value 3 by two numbers, namely, 3.11 and 2.92.
4. From column (1), select the sample size $n = 9$ which corresponds to the smaller of the bracketing values.

5. From column (2) obtain the corresponding $W_a = 5.08$ to determine $R_a = W_a \cdot \sigma'_{AQL} = 5.08 \times 5 = 25.40$.

To summarize the rule:

1. If $\sigma'_{AQL} = 5$, $\sigma'_{RQL} = 15$, $\alpha = .01$, $\beta = .05$
2. Take 9 observations.
3. Calculate the sample range

$$R = x_{max} - x_{min}$$

4. If $R \leq 25.40$, accept the lot, and
   if $R > 25.40$, reject the lot.

To determine the OC function one has to consider $P(R \leq R_a | \sigma') = P_a$, the probability of accepting a lot given that σ' is the true standard deviation. This time we look to the probabilities associated with the distribution of the range.

$$\text{Now } P(R \leq R_a | \sigma') = P\left(\frac{R}{\sigma'} \leq \frac{R_a}{\sigma'} \middle| \sigma'\right) = P_a$$

But for a given σ', the distribution $\frac{R}{\sigma'}$ is given as a probability integral of the range $W = \frac{R}{\sigma'}$ in samples of size n. Therefore, we need to use only

$$P\left(W \leq \frac{R_a}{\sigma'} \middle| \sigma' \right) = P_a.$$

We substitute several values of σ' to determine the protection.[5]

TABLE II: Values for the OC Curve
Single Sampling Plan for Variability
$\sigma'_{AQL} = 5$, $\sigma'_{RQL} = 15$, $\alpha = .01$, $\beta = .05$;
Plan: $n = 9$; $R \leq 25.40$ accept, $R > 25.40$ reject.

σ'	$\frac{R_a}{\sigma'}$	P_a
0	∞	100
3	8.47	100
4	6.35	100
5	5.08	99
6	4.23	93
7	3.63	80
8	3.18	63
9	2.82	45
10	2.54	32
12	2.12	14
15	1.69	4
20	1.27	0.7

FOURTH METHOD OF SOLUTION: Unit Sequential Analysis for Variability.[6]

Unit sequential sampling exists when observations are taken in sequence, At each stage of sampling, one has the opportunity to make one of three decisions, namely, (1) Accept the lot, (2) Reject the lot, (3) Continue sampling by taking another observation.

It is "well known" that a sequential sampling procedure is uniquely determined if we have the four quantities which we used to determine the single sampling plan. That is, σ'_{AQL}, σ'_{RQL}, α, β. The most interesting nuance with respect to unit sequential sampling is that the n is a random variable. One does not choose a fixed value of n before sampling. There are three considerations which represent the characteristics of the sequential plan. These are

1. PLAN - The decision making rule.
2. PROTECTION - The OC function.
3. COST - The ASN function.

THE PLAN:

There are just three numbers that we must determine from the quantities σ'_{AQL}, σ'_{RQL}, α, and β. Label these numbers h_1, h_2, and s, where

$$h_1 = \frac{2 \log_e \frac{1-\alpha}{\beta}}{D}$$

$$h_2 = \frac{2 \log_e \frac{1-\beta}{\alpha}}{D}$$

$$s = \frac{\log_e \sigma'^2_{RQL}/\sigma'^2_{AQL}}{D}$$

where $$D = \frac{1}{\sigma'^2_{AQL}} - \frac{1}{\sigma'^2_{RQL}}$$

Our interest is to present a chart representing the unit sequential acceptance PLAN for variability. The only task that we have is to plot two straight lines parallel to each other as in Figure 1. The equations of the lines are

$$L_1 = -h_1 + sn$$

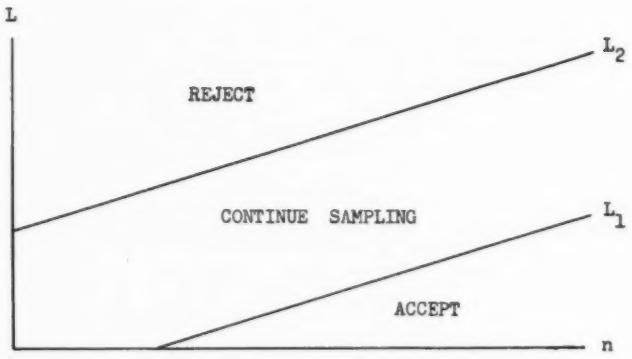
$$L_2 = h_2 + sn$$

where $-h_1$, h_2, and s are respectively the two intercepts and the slope of the lines, n is <u>one less</u> than the number of observations taken sequentially, and

$$L = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n+1}$$

where each of the above summations is over the $n+1$ observed values of x_i.

FIGURE 1:



The sequential procedure is as follows:

1. Draw the two parallel lines as defined by σ'_{AQL}, σ'_{RQL}, α, and β.

2. Label the vertical axis L .

   Label the horizontal axis n.

3. Start taking observations (two the first time) and thereafter adding one observation at a time.

4. After each observation calculate the value of L and plot it against the value of n. (Note: When the mean is unknown, the acceptance and rejection of the lot are based on n which is one less than the total number of observations.)

5. At any stage, if the point falls above the upper line, reject the lot; or below the lower line, accept the lot.

6. As long as the point falls within the two parallel lines, take another observation.

7. Continue steps 3, 4, 5, and 6 until a decision of acceptance or rejection is reached.

The PROTECTION:

Ordinarily, it is not necessary to determine the OC curve by more than five convenient points as follows:

TABLE III: Values of the OC curve
Unit Sequential Sampling Plan for Variability
σ'_{AQL}, σ'_{RQL}, α, β.

σ'	0	σ'_{AQL}	s	σ'_{RQL}	∞
P_a	100	$1-\alpha$	$\frac{h_2}{h_1 + h_2}$	β	0

However, if one wishes to plot more points on the OC curve, following is the set of parametric equations in t which describe P_a for given values of σ':

$$\sigma' = \sqrt{\frac{1 - e^{-2ts}}{2t}} \quad ; \qquad P_a = \frac{e^{t(h_1+h_2)} - e^{th_1}}{e^{t(h_1+h_2)} - 1}$$

See references (3) and (6).

The COST:

The average sample number (ASN) gives the relation between the given standard deviation σ' and the expected number of observations necessary to make a decision. The ordinate of the ASN curve for any given σ' is expressed as

$$(ASN)_{\sigma'} = \frac{P_a (h_1+h_2) - h_2}{s - \sigma'^2} + 1$$

The above formula can be evaluated without embarrassment except for $\sigma' = \sqrt{s}$. Under this situation it can be shown that

$$(ASN)_{\sqrt{s}} = \frac{h_1 h_2}{2s^2} + 1 .$$

This value will ordinarily be close to the maximum average amount of inspection required to reach a decision. However, again the five points, as follows, usually give sufficient information to sketch the ASN curve.

TABLE IV: Values of the ASN Curve
Unit Sequential Sampling Plan for Variability
σ'_{AQL}, σ'_{RQL}, α, β.

σ'	ASN
0	$\frac{h_1}{s} + 1$
σ'_{AQL}	$\frac{(1-\alpha)h_1 - \alpha h_2}{s-\sigma'^2_{AQL}} + 1$
$\sqrt{s}$	$\frac{h_1 h_2}{2s^2} + 1$
σ'_{RQL}	$\frac{h_1 - (1-\beta)h_2}{s-\sigma'^2_{RQL}} + 1$
∞	2

Statement of Another Problem:

Consider the problem in which a set of n similar units are required for a particular assembly. There is the specification that these units do not vary too greatly about any mean value that they might have. The specification might read that the range R of the set of n observations does not exceed a given constant c. That is, $R \leq c$ for a set of n observations. (3)

We know that a process will not make every set of n units such that the range for each set is always less than or equal to a given constant. However, we can state that an acceptable process is one that produces, say, 98% of the sets, such that their range is less than or equal to 0.5 of a unit of measure. For definiteness let us suppose $n = 20$ units per set. We wish a producer's risk of $\alpha = 0.05$. A process that yields only 85% of the sets, such that their ranges are less than or equal to 0.5 of a unit of measure, is considered unacceptable. We wish a consumer's risk of $\beta = .10$.

If we assume that the normal distribution is adequate to represent our distribution of measures, we can find the corresponding values of σ'_{AQL} and σ'_{RQL} associated with the above acceptable and unacceptable quality levels. We know that the sample size in this situation is $n = 20$. Therefore, we can inquire of the tables for the distribution of $W = \frac{R}{\sigma'}$, as to that value of W which satisfies the statement

$$P(R \leq W\sigma' | \sigma') = .02 \quad .(5)$$

We find that this value of W is 5.40. Now substituting $R = .5$ and $W = 5.40$ in the equation $\sigma'_{AQL} = \frac{R}{W} = \frac{.5}{5.40} = .090$. Similarly, $\sigma'_{RQL} = .112$. Now we have the four values

$$\sigma'_{AQL} = .090, \quad \sigma'_{RQL} = .112, \quad \alpha = .05, \quad \beta = .10.$$

It is clear that we can immediately use Table A to determine either of the two values, σ_a or $n\sigma_a^2$ to be used in a single sampling plan for variation. In our case, the decision rule becomes

1. Take one set of 20 units and measure each item.
2. Calculate the sample sum of squares, SS.
3. If $SS \leq n\sigma_a^2$, accept the process making sets of 20 units, and if $SS > n\sigma_a^2$, reject the process.

The OC curve can likewise be obtained as previously explained when discussing single sampling by using the sample standard deviation, σ. It is interesting to note that in this particular problem there is a switch; namely, we are using the standard deviation as a criterion for decision about the range.

TABLE A: Single Sampling for Variability Using the Sample σ or SS.
Given: σ'_{AQL}, σ'_{RQL}, α, β

Sample Size	χ^2_{α}	$\chi^2_{1-\beta}$	$(\sigma'_{RQL}/\sigma'_{AQL})^2$	$\chi^2_{1-\beta}$	$(\sigma'_{RQL}/\sigma'_{AQL})^2$	$\chi^2_{1-\beta}$	$(\sigma'_{RQL}/\sigma'_{AQL})^2$
n	α=.01	β=.01	α=.01 β=.01	β=.05	α=.01 β=.05	β=.10	α=.01 β=.10
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3	9.21	.02	458.21	.10	89.42	.21	43.65
4	11.34	.12	98.65	.35	32.23	.58	19.43
5	13.28	.30	44.70	.71	18.67	1.06	12.48
6	15.09	.55	27.23	1.14	13.18	1.61	9.37
7	16.81	.87	19.28	1.64	10.28	2.20	7.63
8	18.48	1.24	14.91	2.17	8.53	2.83	6.52
9	20.09	1.65	12.21	2.73	7.35	3.49	5.76
10	21.67	2.09	10.38	3.32	6.52	4.17	5.20
11	23.21	2.56	9.07	3.94	5.89	4.86	4.77
12	24.72	3.05	8.10	4.58	5.40	5.58	4.43
13	26.22	3.57	7.34	5.23	5.02	6.30	4.16
14	27.69	4.11	6.74	5.89	4.70	7.04	3.93
15	29.14	4.66	6.25	6.57	4.43	7.79	3.74
16	30.58	5.23	5.85	7.26	4.21	8.55	3.58
17	32.00	5.81	5.51	7.96	4.02	9.31	3.44
18	33.41	6.41	5.21	8.67	3.85	10.08	3.31
19	34.80	7.02	4.96	9.39	3.71	10.86	3.20
20	36.19	7.63	4.74	10.12	3.58	11.65	3.11
21	37.57	8.26	4.55	10.85	3.46	12.44	3.02
22	38.93	8.90	4.38	11.59	3.36	13.24	2.94
23	40.29	9.54	4.22	12.34	3.27	14.04	2.87
24	41.64	10.20	4.08	13.09	3.18	14.85	2.80
25	42.98	10.86	3.96	13.85	3.10	15.66	2.74
26	44.31	11.52	3.85	14.61	3.03	16.47	2.69
27	45.64	12.20	3.74	15.38	2.97	17.29	2.64
28	46.96	12.88	3.65	16.15	2.91	18.11	2.59
29	48.28	13.56	3.56	16.93	2.85	18.94	2.55
30	49.59	14.26	3.48	17.71	2.80	19.77	2.51
31	50.89	14.95	3.40	18.49	2.75	20.60	2.47

TABLE A: Single Sampling for Variability (Continued)
Given: σ'_{AQL}, σ'_{RQL}, α, β

Sample Size	χ^2_{α}	$\chi^2_{1-\beta}$	$(\sigma'_{RQL}/\sigma'_{AQL})^2$	$\chi^2_{1-\beta}$	$(\sigma'_{RQL}/\sigma'_{AQL})^2$	$\chi^2_{1-\beta}$	$(\sigma'_{RQL}/\sigma'_{AQL})^2$
n	α=.05	β=.01	α=.05 β=.01	β=.05	α=.05 β=.05	β=.10	α=.05 β=.10
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3	5.99	.02	298.06	.10	58.16	.21	28.39
4	7.82	.12	67.96	.35	22.20	.58	13.36
5	9.49	.30	31.95	.71	13.34	1.06	8.92
6	11.07	.55	19.98	1.14	9.67	1.61	6.88
7	12.59	.87	14.44	1.64	7.70	2.20	5.71
8	14.07	1.24	11.35	2.17	6.49	2.83	4.97
9	15.51	1.65	9.42	2.73	5.67	3.49	4.44
10	16.92	2.09	8.10	3.32	5.09	4.17	4.06
11	18.31	2.56	7.16	3.94	4.65	4.86	3.76
12	19.68	3.05	6.44	4.58	4.30	5.58	3.53
13	21.03	3.57	5.89	5.22	4.02	6.30	3.34
14	22.36	4.11	5.44	5.89	3.80	7.04	3.18
15	23.68	4.66	5.08	6.57	3.60	7.79	3.04
16	24.99	5.23	4.78	7.26	3.44	8.55	2.92
17	26.30	5.81	4.52	7.96	3.30	9.31	2.82
18	27.59	6.41	4.31	8.67	3.18	10.08	2.74
19	28.87	7.02	4.12	9.39	3.07	10.86	2.66
20	30.14	7.63	3.95	10.12	2.98	11.65	2.59
21	31.41	8.26	3.80	10.85	2.89	12.44	2.52
22	32.67	8.90	3.67	11.59	2.82	13.24	2.47
23	33.92	9.54	3.56	12.34	2.75	14.04	2.42
24	35.17	10.20	3.45	13.09	2.69	14.85	2.37
25	36.42	10.86	3.35	13.85	2.63	15.66	2.33
26	37.65	11.52	3.27	14.61	2.58	16.47	2.29
27	38.88	12.20	3.19	15.38	2.53	17.29	2.25
28	40.11	12.88	3.11	16.15	2.48	18.11	2.21
29	41.34	13.56	3.05	16.93	2.44	18.94	2.18
30	42.56	14.26	2.99	17.71	2.40	19.77	2.18
31	43.77	14.95	2.93	18.49	2.37	20.60	2.13

TABLE A: Single Sampling for Variability (Continued)
Given: σ'_{AQL}, σ'_{RQL}, α, β

Sample Size	χ^2_{α}	$\chi^2_{1-\beta}$	$(\sigma'_{RQL}/\sigma'_{AQL})^2$	$\chi^2_{1-\beta}$	$(\sigma'_{RQL}/\sigma'_{AQL})^2$	$\chi^2_{1-\beta}$	$(\sigma'_{RQL}/\sigma'_{AQL})^2$
n	α=.10	β=.01	α=.10 β=.01	β=.05	α=.10 β=.05	β=.10	α=.10 β=.10
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3	4.60	.02	229.10	.10	44.71	.21	21.82
4	6.25	.12	54.36	.35	17.76	.58	10.70
5	7.78	.30	26.19	.71	10.95	1.06	7.31
6	9.24	.55	16.67	1.14	8.07	1.61	5.74
7	10.64	.87	12.21	1.64	6.51	2.20	4.83
8	12.02	1.24	9.70	2.17	5.55	2.83	4.24
9	13.36	1.65	8.12	2.73	4.89	3.49	3.83
10	14.68	2.09	7.03	3.32	4.42	4.17	3.52
11	15.99	2.56	6.25	3.94	4.06	4.86	3.29
12	17.28	3.05	5.66	4.58	3.78	5.58	3.10
13	18.55	3.57	5.19	5.23	3.55	6.30	2.94
14	19.81	4.11	4.82	5.89	3.36	7.04	2.81
15	21.06	4.66	4.52	6.57	3.21	7.79	2.70
16	22.30	5.23	4.27	7.26	3.07	8.55	2.61
17	23.54	5.81	4.06	7.96	2.96	9.31	2.53
18	24.77	6.41	3.87	8.67	2.86	10.08	2.46
19	25.99	7.02	3.70	9.39	2.77	10.86	2.39
20	27.20	7.63	3.56	10.12	2.69	11.65	2.33
21	28.41	8.26	3.44	10.85	2.62	12.44	2.28
22	29.62	8.90	3.33	11.59	2.56	13.24	2.24
23	30.81	9.54	3.23	12.34	2.50	14.04	2.19
24	32.01	10.20	3.14	13.09	2.44	14.85	2.16
25	33.20	10.86	3.06	13.85	2.40	15.66	2.12
26	34.38	11.52	2.98	14.61	2.35	16.47	2.09
27	35.56	12.20	2.92	15.38	2.31	17.29	2.06
28	36.74	12.88	2.85	16.15	2.27	18.11	2.03
29	37.92	13.56	2.80	16.93	2.24	18.94	2.00
30	39.09	14.26	2.74	17.71	2.21	19.77	1.98
31	40.26	14.95	2.69	18.49	2.18	20.60	1.95

TABLE B: Single Sampling for Variability Using the Sample Range R.
Given: σ'_{AQL}, σ'_{RQL}, α, β

Sample Size	$W_{1-\alpha}$	W_β	$\sigma'_{RQL}/\sigma'_{AQL}$	W_β	$\sigma'_{RQL}/\sigma'_{AQL}$
n	α=.01	β=.01	α=.01 β=.01	β=.05	α=.01 β=.05
(1)	(2)	(3)	(4)	(5)	(6)
2	3.64	.02	182.00	.09	40.44
3	4.10	.22	18.64	.45	9.11
4	4.38	.47	9.32	.77	5.69
5	4.59	.70	6.56	1.04	4.41
6	4.74	.89	5.32	1.26	3.76
7	4.87	1.07	4.55	1.44	3.38
8	4.98	1.22	4.08	1.60	3.11
9	5.08	1.36	3.73	1.74	2.92
10	5.15	1.48	3.48	1.86	2.77
11	5.22	1.59	3.28	1.97	2.65
12	5.28	1.69	3.12	2.07	2.55

Sample Size	$W_{1-\alpha}$	W_β	$\sigma'_{RQL}/\sigma'_{AQL}$	W_β	$\sigma'_{RQL}/\sigma'_{AQL}$
n	α=.05	β=.01	α=.05 β=.01	β=.05	α=.05 β=.05
(1)	(2)	(3)	(4)	(5)	(6)
2	2.77	.02	138.50	.09	30.78
3	3.34	.22	15.18	.45	7.42
4	3.65	.47	7.77	.77	4.74
5	3.87	.70	5.53	1.04	3.72
6	4.04	.89	4.54	1.26	3.21
7	4.18	1.07	3.91	1.44	2.90
8	4.29	1.22	3.52	1.60	2.68
9	4.39	1.36	3.23	1.74	2.52
10	4.48	1.48	3.03	1.86	2.41
11	4.55	1.59	2.86	1.97	2.31
12	4.62	1.69	2.73	2.07	2.23

References

1. Burr, Irving W., "Engineering Statistics and Quality Control", McGraw-Hill Book Co., Inc., 1953.

2. Duncan, Acheson J., "Quality Control and Industrial Statistics", Richard Irwin, Inc., Homewood, Ill, 1955.

3. Eisenhart, Hastay, and Wallis, "Techniques of Statistical Analysis", McGraw-Hill Book Co., Inc., 1947.

4. Pearson, E. S., "The Percentage Limits for the Distribution of Range in Samples from a Normal Population", Biometrika, Vol. XXIV (1932) pp. 404-17.

5. Pearson, E. S., "The Probability Integral of the Range in Samples of N Observations from a Normal Population", Biometrika, Vol. XXXII (1941-42) pp. 301-308.

6. Wald, A., "Sequential Analysis", John Wiley and Sons, Inc., 1947.

SAMPLING INSPECTION BY VARIABLES IV: AVAILABLE SAMPLING PLANS

R. L. Storer
Phillips Petroleum Company

Suppose that one has examined the variables inspection technique and is persuaded that it offers some possibility of use in a particular inspection problem. What comes next? Is it necessary to create a plan to suit the problem? What other factors must be considered prior to installing a plan for a trial run?

Building a plan from scratch will probably be unnecessary. Four good sets of plans are readily available from which to choose. At least one of the two or three hundred plans in these sets should meet the need with little, if any, modification.

As for "other factors", it is suggested only that the prospective user take a last, close look at the question whether a variables inspection plan is appropriate to his needs. In particular, he should consider the matters of cost, purpose of the inspection, distribution of the quality characteristic which interests him, and ease of administration.

Some of the earliest development of variables inspection for percent defective in this country was occasioned by situations in which cost was a secondary consideration. The problem was to make the most of a sample which, by any standards, was too small. The product was expensive and complicated, the total number manufactured small, the inspection totally destructive. Variables inspection was a means of extracting maximum information from an inspection which would be extremely costly regardless of how the data were used.

Most industrial applications are, unfortunately, less clean-cut. Given a certain number of items, M, on which decisions must be made, the costs involved can be represented as follows:

$F = fM$, the total number of pieces to be inspected if the fraction inspected is f.

$C = Fc$, the total cost of inspecting the material, neglecting rejections, when the cost of inspecting one unit is c.

E, the cost of inspection equipment required.

$D = Ld$, the cost of reaching decisions on the inspection data when the number of decisions (lots) is L and the cost per lot is d.

$R = Fr$, the total cost of product destroyed or damaged by inspection when r is the unit cost of the product, including cost of repair where applicable.

$W = C + E + D + R$ is the total cost of reaching the decisions.

The cost elements, obviously, are rather more complicated than a simple determination as to which method requires the more inspection. For simple inspections on small samples, the indications are that the attri-

butes approach enjoys some cost advantage. As the sample size increases the inspection cost differential tends to decrease, though still favoring the attribute approach to some extent, but the decision (computation) costs increase. In more complicated dimensional inspections, the situation may well be reversed. Inspection setups for such characteristics as the alignment of several points, or the relative position of several surfaces, may be more expensive both to instrument and to operate on an attributes basis than on a variables basis. In any case, inspectors tend to regard the variables approach as "more trouble". It is well to make a fairly complete cost analysis in order to be in position to give a definitive answer to this charge, if for no other reason. When the required figures are unavailable, catalogs, a table of standard times, a stopwatch and a few hours should yield estimates good enough for comparative purposes.

Purpose is often the determining factor in the decision between the attributes and variables approaches. It is one thing to select a plan for deciding whether to buy or not to buy a lot of product. Quite another thing to select a plan appropriate to a decision whether a lot of material is suitable for a certain contemplated use, and something else again if the purpose is control of production. Sometimes decisions are required for a project which is a mixture of these purposes. New products, for example, usually demand more attention and information than products which have been satisfactorily produced over a period of months or years. A variables plan might well be indicated for such an application, regardless of any real or supposed advantages of the attributes approach. Many other engineering situations exist where the extra costs of measurement and computation, even when added to a sample large enough for an attributes decision, seem worthwhile. Discussion with the product designers or users should be helpful in resolving this point.

Note that the sampling plans under discussion here are primarily aimed at the situation where interest lies in fraction defective, or joint variation of mean and dispersion. The same data may be used where interest centers primarily on the mean or the disjoint variation of mean and dispersion, or for control charting. But the decision procedure $\bar{x} \pm ks \lesssim U, L$ is not appropriate to such applications.

Sometimes the nature of the inspection or test determines whether attributes or variables will be used. Tensile strength determinations, for example, are seldom reported on an attribute basis because of the way in which they are derived. Since the reading must be obtained from a graduated dial in any case, it is sensible to use it as reported rather than convert it to "good" or "bad". If you are not familiar with the inspection operation, or know it only by reports from others, personel observation would be a good idea prior to disposing this factor.

A factor which can cause unnecessary trouble is the form of distribution of the characteristic inspected. One hears that detailed evidence of normality is necessary to the successful use of variables inspection -- and on the other hand that the type of distribution makes no difference. If the first be true, then so much must be known about the product in advance of inspection that the necessity for inspection is questionable. If the latter be true, then many excellent men have spent much time trying to solve a non-existent problem.

As usual, the truth is somewhere between the two extremes. If a

fairly substantial proportion of product exceeding the limits U and L has been removed from the material by screening prior to sampling, someone is going to be most unhappy with the rejection decision which is almost sure to follow in the wake of application of a simple, unmodified variables plan. The effect is the same as if the plans possessed a sixth sense that non-conforming product was once in the lot, and lacked the horse sense to recognize that it had been removed. An attributes plan, or a combined variables-attributes scheme of some sort is indicated for such a situation.

Slight departures from the normal curve model will cause trouble only if both persistent (rather than intermittent) and unrecognized. If it is found that a distribution is "skewed" rather than symmetrical, the multiplier "3" which is customarily used on sigma to obtain end-points of the distribution must be changed to "2" or "4" or "5" or some other number. Knowing this, we can either use plans based upon the multiplier "3" and adopt a different or more liberal interpretation of an indication that a limit has been exceeded; or, we may concoct a new plan based upon a multiplier appropriate to the distribution model which describes our problem more exactly. A Control Chart test on some preliminary data, or any of several "quick and dirty" tests for normality should give the desired information rather quickly.

One last point to be considered in the "whether variables" decision: Ease of administration. It is easier to teach an inspector to use a snap gage correctly than to teach him to use a dial gage correctly. He is more likely to be consistently accurate with the snap gage than the dial. Attribute reporting forms are generally simpler and less subject to arithmetical error than variables reporting forms. At least, this has been my experience, and I find that many inspection administrators agree. These factors must be carefully weighed where a substantial number of inspectors is involved or the inspection locations are so widely dispersed that supervision is difficult.

Assuming that this "last look" at suitability does not change the original assessment that a variables plan will fill the bill, a scanning of the available sets of plans is in order.

The earliest widely available set of tables was published by Bowker and Goode [1] in 1952. Their book contains, in addition to the tables, a thorough and easily read exploration of the development and use of variables inspection plans for the fraction defective.

Their tables utilize only the sample standard deviation as a measure of variability, but are otherwise the most comprehensive set available at time of writing. Fifteen AQL classes and fourteen AOQL classes are tabulated. Separate sets of plans are given for the situation where the standard deviation is unknown, and for the situation where some previous history is available for its estimation. k values in the former set are somewhat smaller than in the latter, leading to more conservative decisions.

The tabulated plans are for use against one-sided specifications: Use of two one-sided plans is recommended for the situation where a history is available from which to estimate the standard deviation and the difference U-L is at least M times the standard deviation. M is a tabulated factor which varies with the AQL. When $U - L < M\hat{\sigma}$, a graphical decision procedure is provided, the acceptance region for the

graph being constructed with the assistance of a table of factors classified by sample size and AQL.

Operating characteristic curves are provided for plans operating against one-sided limits. Normal, reduced and tightened inspection are available as well as single and double sampling.

A table of control chart factors is provided for use with standard deviations computed on n-1 rather than n.

The plans have been constructed in accordance with the concept that the AQL is that percent defective for which the probability of acceptance is about 95%, regardless of lot size. Operating characteristics were determined with a view to companionate use of the attribute plans constructed for the Navy by the Columbia Research Group.[2] This pair of books is an excellent base for a systematic inspection scheme of general coverage. Since both AQL and AOQL classifications are available, a set of Continuous Sampling Inspection Plans can be easily tied in, providing plans for accept/reject decisions in a wide variety of production and inspection situations.

Navy Bureau of Ordnance tables for inspection by variables[3] were issued in 1952 as a government document rather than through commercial channels. Comparison to the Bowker and Goode Tables reveals several differences and some similarities.

All the Navy plans utilize standard deviation as a measure of variability. Six AQL classes are provided. No classification by AOQL is offered. Normal, tightened and reduced inspection are available, but double-sampling plans have been omitted.

Each plan is designed to have operating characteristics similar to the MIL-STD-105A plan (attributes) of equivalent AQL and lot size. The probability of accepting material which is AQL percent defective therefore increases as the lot size increases, rather than remaining at 95% for all lot sizes. Every effort has been made to make OC curves match those in MIL-STD-105A as closely as possible.

The outstanding feature of these plans, as compared to others presently available, is the form in which the inspection data is processed to reach a decision. Bowker and Goode, and the Army Ordnance tables are both based upon finding an average, $\bar{x}$, to which is added (or subtracted) a factor obtained by multiplying the observed standard deviation, s, by a factor k. The result, for example $\bar{x} + ks$, is then directly compared to a given limit for the characteristic, say, U.

The Navy tables handle this procedure on the basis of a "Lot Quality Index", the derivation of which is as follows. Suppose $\bar{x} + ks = U$. Then $k = \frac{U - \bar{x}}{s}$. If we add 10 to both sides of this equation, and give the quantity $k + 10$ the designation C_U, we obtain the Lot Quality Index $C_U = \frac{U - \bar{x}}{s} + 10$. In practice, C_U (or C_L) is computed from the inspection data and compared to the tabulated value of the index, called A_U (or A_L) for the particular AQL and sample size being used, the decision being based upon whether the tabular value is, or is not, exceeded.

Another innovation is the Maximum Allowable Standard Deviation

(MASD) concept for the case where inspection against a two-sided limit is required. MASD is defined and computed in the form $\frac{U-L}{C_U - C_L}$. The results must meet all of the criteria $C_U \geq A_U$, $C_L \leq A_L$, $s \leq$ MASD. This procedure avoids the necessity for a special graphical solution, as required by the Bowker & Goode plans. A worthwhile practical advance, since most inspectors for some reason share an aversion to this kind of graphical decision. It has the added advantage of protecting against the situation where the standard deviation is too large for the safe use of two one-sided plans, but without introducing the idea of "known standard deviation" which is confusing to some people.

The inspection instructions included in these tables are naturally less likely to be suitable for general industrial use than the Bowker and Goode plans because of the more specialized use intended for the Navy document. However, the tables themselves are compact and quite useful in many industrial situations. A standard computation form is provided which is very efficient for any procedure which requires computation of the standard deviation.

The Army Ordnance tables were published in 1954. They use the constant-risk principle employed by Bowker and Goode; the probability of acceptance for material of AQL quality is 95%, regardless of lot size. Eleven AQL classes are provided; AOQL classification is omitted, in accordance with recent government practice.

The features which distinguish these tables from the other two are as follows. A set of plans is provided which utilizes the range as a measure of variation, in addition to the usual set utilizing the standard deviation. The tables are compact. A single page serves to express all the information (except OC curves) required to use either the range or the standard deviation plans. The MASD concept has been simplified to a factor, F, which is listed in the body of the tables along with the constant, k.

The two sets of plans (range and standard deviation) have been so arranged that the protection afforded by a given sample size and AQL is about the same, regardless of which measure of variation is chosen. The standard deviation plans accept less material of worse-than-AQL quality than do the range plans, but the difference is so slight that a single set of Operating Characteristic Curves suffices to describe both sets of plans, for practical purposes.

As with the Bowker and Goode plans, decision is made in terms of data as observed, without conversion to intermediate factors or fraction defective. In the two-sided case, which is the most complicated, the three criteria required are: $\bar{x} + kV \leq U$, $\bar{x} - kV \geq L$, $V \leq F\,(U - L)$.

In 1955 Lieberman and Resnikoff published (7) a new set of sampling tables developed at Stanford, along with a summary of the considerable theoretical work which led to these plans. These tables afford a choice among fourteen AQL classes. Unknown standard deviation, known standard deviation and average range plans are provided. Single sampling at one level of inspection is used throughout but reduced inspection is in effect available by switching from "unknown" to "known" standard deviation when sufficient data has been collected. Operating characteristic curves are available for all plans, including those for use against

two-sided limits.

AQLs follow the MIL-STD-105A practice, the probability of acceptance for AQL quality ranging from about .89 for the smallest sample size to about .99 for the largest sample size. A single OC curve describes unknown standard deviation, known standard deviation or average range plans for a given AQL against either a one-sided or a two-sided limit. However the sample size is given by code letter and will be found to differ with the measure of variation chosen. Code letter "I", for example, represents a sample size of 5 if a known standard deviation is used; 25 for unknown standard deviation; 30 for an average range plan.

Decisions are reached by a three stage process. First, the quantity $C = \frac{U - \bar{x}}{s}$, or its equivalent for lower limits and other measures of variation, is computed. Then C is used to enter a table which gives an estimate $\bar{p}$ of the fraction defective of the lot from which the sample was drawn. $\bar{p}$ is then compared to a value p^* obtained from the sampling table.

If $\bar{p} \leq p^*$, accept; otherwise, reject. When a two sided specification is involved, the estimates for the fraction beyond upper and lower limits are added and compared to p^*, that is, $P_U + P_L \leq p^*$ is the criterion of acceptance.

As of the time of this writing, a Military Standard for Inspection by Variables has been in preparation for something like three years. Although it is nearing the stage of publication, it is not yet ready. In any case, information concerning its content should properly come from a government employee. I hope, however, that when it does appear it will be found to combine the best features of existing sets of plans in such a way as to provide a document of widespread usefulness in both industry and government.

If none of the published sets of plans seem to fit your requirements, References 1, 2, and 5 through 11 provide the theory and some computational suggestions for constructing additional plans.

Selection of an appropriate plan from a given set boils down to decisions on measure of variation, AQL and sample size. The relative efficiency of the range and standard deviation need not enter into the selection since equivalent AQLs provide equivalent operating characteristics regardless of the statistic chosen. In the Army tables the choice of statistic makes no difference in sample size. In the Lieberman and Resnikoff tables it makes a considerable difference. From a practical standpoint, the range is the preferred measure for routine accept/reject inspections. Standard deviation is usually chosen for inspections having an engineering use.

I would like to include a short set of rules for choosing an appropriate AQL, but so far as I know these rules have not yet been devised. Generally, I would assign the smaller AQLs for characteristics of greater importance, larger AQLs for those of lesser importance. If the distribution has a tendency toward long tails on one or both sides, the AQL may have to be made somewhat larger than usual if the specification limits are relatively narrow. Or, it may be necessary to open up the limits somewhat. If this is not done, the plan will reject more frequently than anticipated from the AQL, assuming that AQL quality material is

made and presented for inspection.

In each of the tables mentioned except the Lieberman and Resnikoff tables, sample size is fixed by determining an inspection level and a lot size. Where destructive tests are concerned, it may be advisable to disengage this automatic selection feature and pick a constant sample size which management can or will pay for. This is frequently much smaller than the tables require. The Army tables provide a "category of inspection" label to facilitate such handling and to remove any doubt which might exist on the question whether such a departure is proper. Wherever possible, I recommend citing the sample size and acceptance constants in the inspection instructions, rather than referring the inspector to the tables. This practice minimizes both the number of documents and the number of ways in which the inspector can make a mistake.

If the circumstances permit, a shakedown trial of the selected plan is an excellent idea. Like any good tool, variables inspection works best in the hands of one experienced in its use.

BIBLIOGRAPHY

1. A. H. Bowker & H. P. Goode, Sampling Inspection by Variables, McGraw-Hill, 1952.

2. Eisenhart, Hastay, Wallis, Techniques of Statistical Analysis, McGraw-Hill, 1947.

3. Sampling Procedures and Tables for Inspection by Variables, Naval Ordnance Standard, NAVORD OSTD 80, May, 1952.

4. Sampling Inspection by Variables, Army Ordnance Inspection Handbook, ORD-M608-10, June, 1954.

5. A. Golub, Designing of Single Sampling Inspection Plans by Variables, Ballistic Research Laboratories Memorandum Report No. 718, September, 1953.

6. I. R. Savage, Mixed Variables and Attributes Plans: The Exponential Case, Stanford University Applied Mathematics and Statistics Laboratory, Technical Report No. 23, June, 1955.

7. G. J. Lieberman and G. J. Resnikoff, "Sampling Plans for Inspection by Variables", Journal of the American Statistical Association, June, 1955.

8. Ireson, W. G. Sampling Tables for Inspection by Variables, Stanford University Applied Mathematics and Statistics Laboratory, Technical Report No. 7, 1952.

9. Ireson, W. G. & Resnikoff, G. J., Sampling Tables for Inspection by Variables Based on the Average Range, Stanford University Applied Mathematics and Statistics Laboratory, Technical Report No. 11, 1952.

10. Resnikoff, G. J., A New Two Sided Acceptance Region for Sampling by Variables, Standard University Applied Mathematics and Statistics Laboratory, Technical Report No. 8, 1952.

11. Resnikoff, G. J., The Distribution of Average Ranges for Subgroups of Five, Standard University, Applied Mathematics and Statistics Laboratory, Technical Report No. 15, 1954.

QUALITY CONTROL IN COMPLEX SYSTEMS

F. J. Sindelar
H. J. Wilkinson
International Business Machines Corporation

Today's quality function finds itself facing a tremendous challenge. Regardless of product, the impacts of new materials, automation capabilities and increasing product complexity on the manufacturing effort it is assisting, coupled with the performance and reliability requirements becoming associated with product end use, are presenting an environment requiring technical and administrative planning of the highest caliber. In the design of such a quality function, serious thought must be given to a system concept of relating part, process and assembly data to the product system. This type of planning must be in addition to the corrective action feedback normally associated with any data taking.

We will attempt to show the evolution of the areas associated with the quality function and explore two examples of the approach to system design in quality control.

The quality function has usually been aligned with these classical functions:

Receiving
Manufacturing
Assembly

However, while the scope of each of these functions has been expanded during the past several years, there has been a tendency toward the generation of spheres of activity due possibly to the specialized product and process techniques involved. This philosophy has obviously tended to isolate these functions with an accompanying failure of feedthrough of information to the area charged with final product evaluation. One might indicate the changes in scope as follows:

I. Receiving Inspection

Raw Material Control
Vendor Rating
Vendor Certification
Component and Sub-assembly Receiving
Life Test Data

II. Manufacturing

Process Control -- Metal Working, Heat Treat, Plating
Gage Control
Tool Inspection
Machine Tool Capabilities

BASIC ORGANIZATION OF A QUALITY FUNCTION

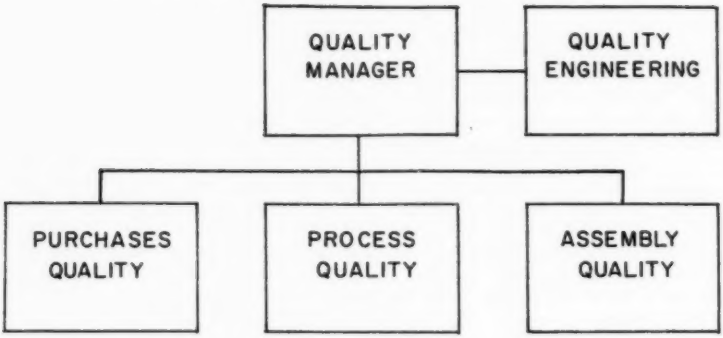


FIG.#1

III. Assembly

Process Control - Electronic Component Manufacture
Life Test Data
Assembly - Product Evaluation
Test Equipment Capabilities
Product Field Performance

Obviously the grouping of the above areas is dependent on the particular company and product involved. The important point at this time is the fact that the areas of involvement have increased considerably. We now realize that we can no longer be satisfied with these sources as individual areas of control. They must necessarily be related to each other as part of a system.

We are at the crossroads in the sense that while Quality Control has been able to show the feasibility of the statistical approach as a practical means of controlling part and process quality we must, in addition, devise ways of combining these data from many areas and relating them to final product evaluation. This must be accomplished in order that the need for costly and often misleading performance testing of complex products might be more thoughtfully considered, interpreted, and thereby reduced or eliminated.

The fact that a given device meets a functional or acceptance test is not necessarily indicative of its reliability in a customer's installation. This knowledge can only come of a complete history consisting of data representing raw material, parts, sub-assemblies, final assembly, and the test experience relating to the unit or system under study.

Figure 1 shows the basic Quality Control Organization, the usual line supporting functions, and with Quality Engineering as a staff function.

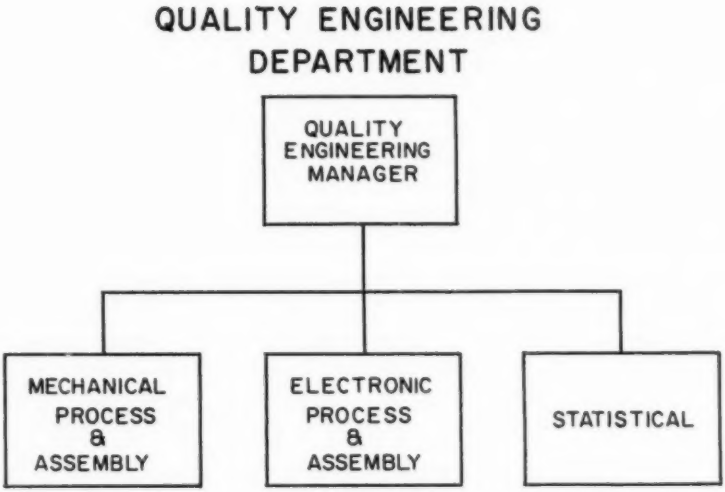


FIG. #2

The concept of a quality engineering group designed to act as a service organization in the quality function is certainly not a new one. However, we would propose that organized system design is not possible without the services of such a group.

Figure 2 shows a basic Quality Engineering Department with individual areas outlined by function. We do not propose to go into great detail regarding this organization since considerable material has already been written on this subject. However, from a system standpoint, we may ask the question, "What can such a group contribute to system design?" We might well review the tools that such a group would of necessity avail itself of in developing system concepts:

A. Knowledge of proven statistical tools
B. Ability to initiate and interpret designed experiments
C. Ability to establish test equipment requirements
D. Experience in and knowledge of manufacturing processes and techniques

A quality engineering group so trained and assigned can give the quality function the type of system thinking which is required in order to completely integrate the incoming materials and the outgoing final product. With such a group assigned to evaluate product quality, management is in a position to probe new areas and new concepts by means of which system thinking can be applied to both product quality and those problems of quality administration so necessary to both the field and factory organizations.

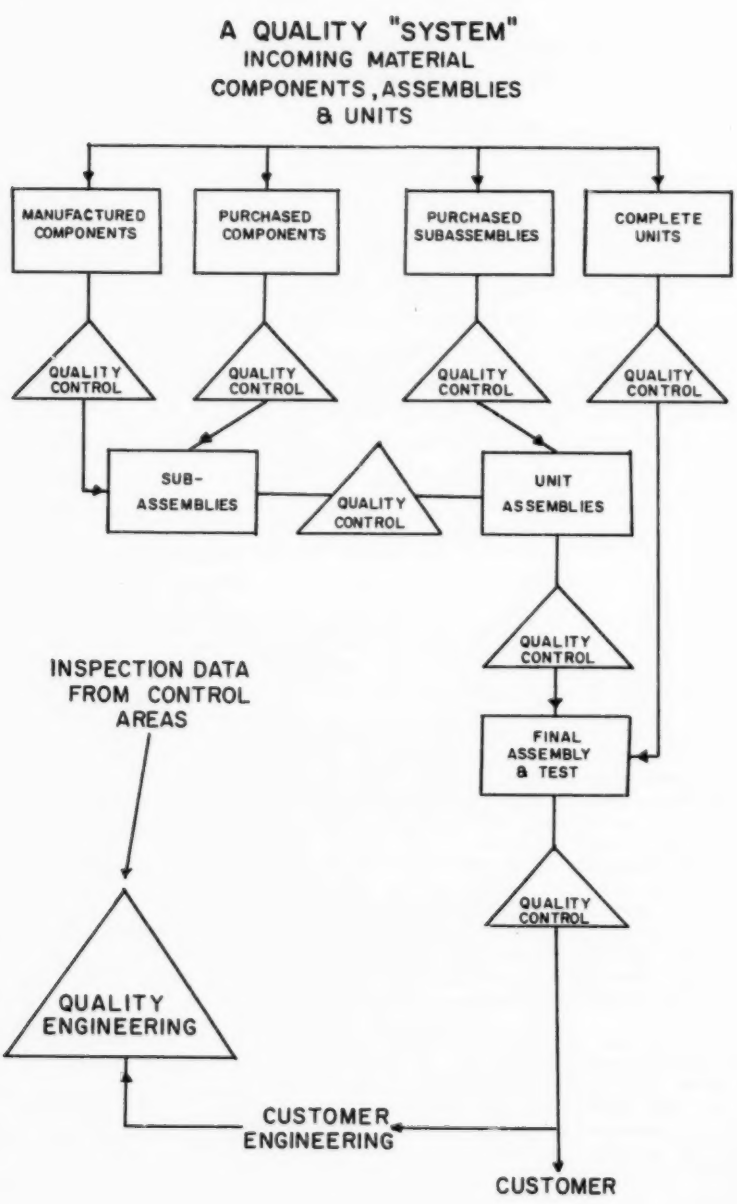

A QUALITY "SYSTEM"
INCOMING MATERIAL
COMPONENTS, ASSEMBLIES
& UNITS
MANUFACTURED COMPONENTS
PURCHASED COMPONENTS
PURCHASED SUBASSEMBLIES
COMPLETE UNITS
QUALITY CONTROL
QUALITY CONTROL
QUALITY CONTROL
QUALITY CONTROL
SUB-ASSEMBLIES
QUALITY CONTROL
UNIT ASSEMBLIES
QUALITY CONTROL
INSPECTION DATA FROM CONTROL AREAS
FINAL ASSEMBLY & TEST
QUALITY CONTROL
QUALITY ENGINEERING
CUSTOMER ENGINEERING
CUSTOMER

FIG. #3

It should be noted that the successful application of both Manufacturing and Quality Engineering analysis must be prefaced by completely adequate Product Engineering specifications.

Figure 3 shows some basic inter-relations of a quality system combining vendor, process, assembly and field information. While in each area we see mechanisms to provide corrective action data for manufacturing within each small system, final product evaluation can be affected to a very marked degree when decisions are based on data gathered progressively from process input to product line output.

A good case and point is the control of a purely manual manufacturing process -- soldering of electrical connections by hand methods. We have here a relatively small sub-assembly subject to a large number of variables, any one of which may seriously affect its all important electrical characteristics.

The following are some of the more obvious variables which we have to contend with in this type of connection:

Surface Contamination
Surface Preparation
Mechanical Stability
Solder Alloys
Fluxes
Temperature
Operator Techniques

The presence of such a critical process in large quantities, as is the case in many phases of communication computer and guidance systems, can have a tremendous effect on product performance and reliability. This can be of particular importance in circuitry involving high frequency pulse logic.

Process evaluation must occur early in product life and must be sustained throughout the assembly and test stages so the final appraisal can be accomplished by audit mechanisms rather than by questionable screening or inconclusive functional testing. Figure 4 presents a system concerned with a manufacturing area receiving three basic types of material. This includes a unit vendor and panel vendor, each building a product to print requirements, and an electronic component vendor supplying an off-the-shelf item.

In the case of the two vendors supplying electronic equipment, acceptance procedures, as noted by the small triangles, constitute the first step in the control and analysis of solder quality. The same acceptance procedures in the Receiving Inspection Department constitute the tool wherein auditing is accomplished. In addition, an analysis procedure is also used in this area in order to audit component surface solderability. Thus we see that the unit and system assembly area is presented with material whose quality levels, with regard to solder quality, are known and under control.

The unit and system assembly area has within it an additional acceptance procedure designed to control and audit any solder

SOLDER CONNECTION QUALITY ANALYSIS

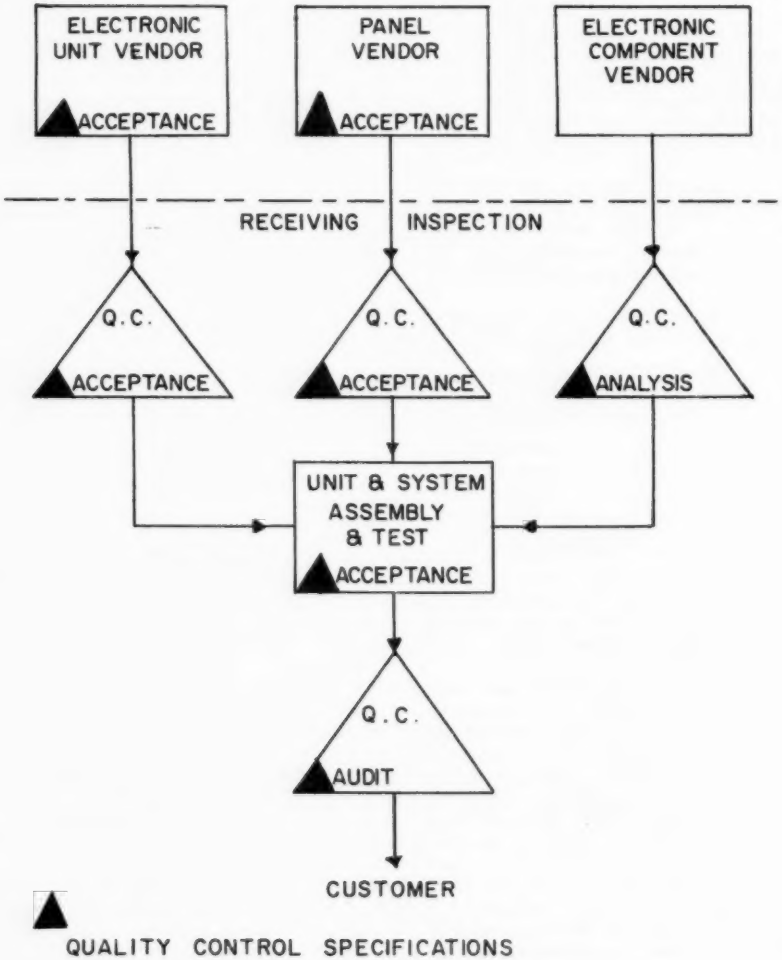


FIG. #4

connections which have been the result of engineering change or repair activity. This type of control system placed the manufacturing group in a position wherein its final product may be audited to determine whether or not the levels previously established have been altered. It should be noted that while the various areas shown here have within them the necessary acceptance, analysis or audit procedures, auxiliary specifications or controls are certainly required for the complete accomplishment of the system philosophy. These may well include the following:

In-plant operator training in solder techniques
Assistance to vendors in operator training
Control of soldering tools
Solder material certification

A second approach to the system concept is shown in Figure 5. This example concerns itself with electron tube quality and reliability analysis. We see here a product controlled by the usual engineering specification and in addition a Quality Acceptance Procedure. This procedure in effect at both the vendors' plants and the receiving inspection area establishes compatible tests and quality levels. In this way, the first check point in product quality is established. The second check point is established by means of a documented line reject program. Data in this area provides a means for reviewing vendor and receiving compatibility, and also provides the first step in the correlation of tube specifications and circuit requirements. The third point in the data system is provided by means of the analysis material made available from the field installations. By proper study, data from these three areas can contribute considerably to the following:

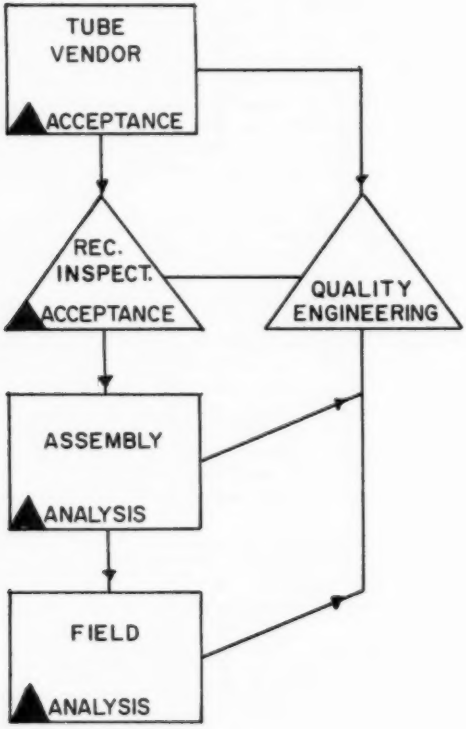
A. Life characteristics
B. Test equipment correlation
C. Tube application
D. Diagnostic techniques

In the course of this discussion we have touched on several different kinds of quality procedures. These documents, under engineering control, are the culmination of the quality engineer's evaluation of the variables and controls which he feels will be associated with a given process. The kinds of procedures will certainly vary dependent on the activity. The following are a few possibilities:

Acceptance
Process Control
Analysis
Audit
Final Test
Reliability
Maintenance

A procedural guide assures a favorable and similar format for procedures in general and also maintains compatibility between several documents of a data gathering system.

ELECTRON TUBE
QUALITY & RELIABILITY
ANALYSIS



▲ QUALITY CONTROL SPEFICATIONS

FIG. #5

The following is a brief table of contents for such procedures:

1.0 General Instructions

The purpose and specific area of coverage is spelled out in this paragraph.

2.0 Classifications of Tests

A general breakdown of tests into specific groups, based on their function in regards to the product, and amount of inspection necessary.

3.0 Sampling Procedure

Instructions on identifying each unit tested, methods of selecting each sample unit, and order in which tests should be made.

4.0 Identification and Handling of Lots

Instruction as to lot size, lot identification, lot disposition (accept or reject) after each test.

5.0 Disposition of Material

Instruction regarding routing of material that has been accepted or rejected.

6.0 Records

A sample of inspection records used with a detailed explanation of their use specify data to be taken and by whom, i.e., manufacturing or inspection.

7.0 Tables and Charts

All tables and charts necessary for this particular procedure are spelled out in detail.

The preceding material has necessarily been brief and probably over simplified as a result. We do not wish to convey the thought that a system approach is simple, nor would we presume that this presentation represents completely acceptable system operation.

However, we have found that rudimentary as they are the system analysis so far installed by us has very measurably enhanced the effectiveness of Quality Control's assistance to our Manufacturing division.

It is very interesting to note that the tools associated with operations research are in essence identical to those currently being used by quality engineering divisions.

The following are some of the accepted techniques:

Flow Charts
Measurements
Experimentation
Information Theory, Boolian Algebra, Gaming, etc.
Simulation
Sampling
Cost Accounting

This is very reassuring in the sense that it is indicative of the evolution of a quality function from the relatively narrow confines of parts and processes to comparitively broad system concepts in both the technical and management aspects of today's manufacturing universe.

USE OF MANAGEMENT REPORTS IN A Q.C. PROGRAM

Harry G. Romig
Summers Gyroscope Company

It is necessary at times to take stock of our Profession. In what areas are we strong? What are our weaknesses? It is apparent that many Managements consider Quality Engineering only an Inspection and Test function, a necessary but undesirable and costly evil that must be subordinated wherever possible. Many Managements must still be shown that, if set up properly to include all its functions, it can be one of its most valuable aids since its functions and activities should cover all phases of the problems that continually harass managements. Simple informative reports and analyses that aid Management in making valid decisions are one part of the varied functions in Quality Engineering that have not received sufficient recognition.

One point should be stressed. As a true Profession, Quality Engineering must become a part of Management, making decisions of its own on the basis of reports and directives from Top Management, such as the Board of Directors, the President and Executive Vice President, to the same extent as Sales, Contracts, Engineering, Production, Accounting, Research and Development, and similar departments or divisions. In turn it provides as one of its functions reports and recommendations as well as services to all other groups as well as Top Management that result in a more economical and efficient operation in all areas related to quality performance. The Customer buys Quality, demands Quality, and must receive Quality materiel in accordance with the provisions of each contract. To provide adequate Assurance and Reliability, one function is to develop and maintain the shortest possible lines of communication and obtain a feedback of information. Management Reports are part of this function.

This presentation covers this important function. It is sub-divided into the following 16 items:

1. Management and Quality Control;
2. A Two-Way Street---Reports to and from Management;
3. Missing Elements;
4. Removing the Rubbish;
5. How Obtain Good Data?;
6. Filling the Gap;
7. How Present Data;
8. Scheduling Reports;
9. What an Administrator Should Know;
10. What a Quality Engineer Should Know;
11. Overlapping Functions---Responsibilities without Authority;
12. Making the Program Tick;
13. Ultimate Goals;
14. Getting into the Act---Making Decisions;
15. Measuring Results; and
16. Maximum Benefits from Management Reports.

1. MANAGEMENT AND QUALITY CONTROL

The administrative aspects in Quality Engineering have been neglected. Quality Engineers have noted this weakness. The Administrative Applications Division of ASQC is its largest division. This presentation has been planned to bring to light some of the reasons Quality Control programs seemingly fade away. It points out how proper use of Management Reports in a Quality Control Program can place the Quality Engineer in his proper Professional status and make this group one of the most valuable in a company in fact, not just in theory.

Impetus was given Quality Control in 1941 and thereafter by war-time needs. These were primarily for Acceptance and Rejection inspections using samples as small as possible to save man-power. Although Quality Control had its beginning in 1924 through the work of Dr. Walter A. Shewhart and his colleagues, only a few of the larger companies had made use of these principles and installed systems. In large part the Military used only two sampling plans: 100% inspection, or 10% inspection, and had no definite criteria for different degrees of seriousness for their 10% samples. Generally the Acceptance Number was zero (0), but was often modified in order to accept production in short supply. Shortage of man-power and increased production demands made it imperative to establish better sampling procedures and to evaluate the quality of products prior to shipment. Levels of performance were established and sampling plans selected that would accept the majority of manufactured parts.

Managements and the Military established Inspection Systems that forced as large a number as possible of workable units out the door each day. Units classed as expendable were built with no consideration for length of useful life or shelf life. Fighter planes were expected to be destroyed early in their life by the enemy, hence were treated as expendable and had only essential performance quality built in. This same pattern was followed for all such so-called expendable units---establish the lowest possible levels of quality in order to maximize production.

The levels of quality for these emergency items were exceedingly low even for this war period. After the war the consumers desired immediate delivery of products of which they had been deprived during the war. In this sellers' market quality had no place. Management's problems primarily were procurement and production. Even for stringent military needs there were no incentives to build quality products that were Reliable and Dependable. Quality Control organizations required by the Military during the war period were disbanded in many instances. This was especially true of those Quality Control engineers not properly trained in true quality control principles and methods. It must be noted, however, that much of this change was due to a shift from war production to commercial production. Many war-time plants were shut down or operated with skeleton crews. The trend was to eliminate

any group that failed to contribute to a company more than its cost. Many Quality Control groups were found wanting.

Concerns with fairly good quality control systems maintained them on a greatly reduced scale, keeping chiefly inspectors and placing most of their other functions in other groups. After the immediate needs of the consumers were satisfied in commercial areas the sellers' market was replaced by a buyers' market. The customers again demanded quality merchandise. Managements had meanwhile added quality functions in other departments. Consequently little quality control existed and practically no quality information or reports. Even the old type of meager Reports on Quality that used to be made to Management and recommendations and Reports from Management were practically non-existent. The Military set up Procurement Specifications with Quality provisions as part of contract requirements, but sometimes these were crossed out in many contracts. Paper-wise adherence to such requirements was usually obtained in some manner. However, field reports indicated that the quality was not as good as expected. As a result the Military began demanding some Assurance of Reliability and often switched contracts from one company to another to obtain maximum performance wherever possible for the price being paid. In many cases contracts were even cancelled.

Currently the demand for Reliability has become so great that Reliability Provisions are often written into the Military contracts. The stress on Quality and Reliability has been so great that many commercial concerns cite their excellent quality and performance in all their advertising media. Both Military and civilian consumers are becoming quality-minded. Managements recognize this and now wish information concerning their quality performance. An impetus has been given to Reliability Programs and also Quality Control, Inspection and Test. Quality Assurance Programs beyond the usual contract and consumer demands have been placed in operation. Managements desire factual reports in order to obtain true measures of their Quality Engineering departments as well as other departments. Also they desire accurate measures of the Expected Quality Performance of their products. It must now be shown the place such Reports have in well-managed companies and how Management may gain most from their use.

2. A TWO-WAY STREET---REPORTS TO AND FROM MANAGEMENT

One of the functions of a Quality Engineering group is to keep Management informed concerning the quality of outgoing product, the nature of the complaints received from customers, the loss due to scrap, the causes of unsatisfactory product, and the corrective measures taken to correct discrepancies of all kinds. It is also necessary to cover the quality of products received from vendors and sub-contractors, and provide some rating system for vendors for use by Purchasing. There may be a number of statistical reports covering special projects, evaluations of various methods and designs, and other features that Management may desire on a regular

periodic basis or possibly upon request. In some companies, where there exists no true separate statistical group, Quality Engineering may be called upon to present weekly, monthly, quarterly and yearly reports to various executives of the company for their regular use. An extreme case may be where the individual members of the Board of Directors and particularly the Officers of the Company desire definite reports provided them several days prior to the meeting of the Board of Directors. In one company this regular information was obtained on a routine basis weekly with a summary report prepared prior to each Board meeting. These provided factual information covering not only Quality and Reliability but all functions and phases of the Company useful to the Board in formulating the policies of the Company.

Quality Engineering also requires directives. Hence it requires certain Management Reports in order to foster and develop its own work to the greatest extent. These Management Reports provide regular information covering the policies of the company as they affect the Quality group. Which products have priorities, which ones are soon to be dropped from production? In what field will effort be made to sell new products? Will the standards required be low or extremely strict? One item might sell for $1.00 if made on a production basis so that the large volume made results in very low unit costs. The same item might be produced to sell for $10 or $100 if made more precisely or of better materials, or might be the same item with more expensive surroundings. A good example of this is a watch whose mechanism is of the highest quality possible. The watch mechanism itself may sell for about $20.00. It may go to one area and have a cheap case added so that the selling price is low. The same watch works may go to another dealer, one who specializes in costly cases. Such finished watches may even sell for up to $1000.00 where some cases may be set with precious jewels. The standard for the watch movement may be the same in both cases, or it might be made less stringent for the low-priced watch. The standards for the cases will probably not be the same due to the great difference in costs and the selling prices. Economic quality levels tend to vary with costs as well as with complexity for the products or components under consideration.

The illustration above represents an extreme condition but in electronic equipment and the like the same extremes may exist. One capacitor may require rigid control over the Q factor while another use may be made of the same capacitor where Q is not important. Rigid controls would be maintained for Q in one case and this capacitor would be given one part number, while no controls would be used for Q in the second case and a different piece-part number would be assigned.

Management decides the area in terms of uses and costs in which products will be sold. This information must be given to other groups in order to set proper economical standards for the products as designed. Management may decide to expand factory facilities in order to make certain component parts that have formerly been supplied by outside vendors.

Which component parts can be made most economically within the plant? Because of inter-relationships between the company and many other outside companies, it may be necessary to continue the purchase of many components even though they can be made cheaper by the company. Management must be provided with all the possible information concerning quality, costs, schedules, availability, and stock. The transition from an outside source to company production or the reverse must be scheduled in such a way as not to disrupt the production processes now in operation.

One shop may be apparently losing money consistently through high costs of production and low sales. However, it may be found that some other company policy is responsible for this seeming loss. For example, a concern made a large number of complex items and also water heaters. A constant overhead was charged against all shops. This high overhead was found to be due to research and development programs in other areas and had no connection with the water heaters. The heater's design was frozen. It was very simple to make and cost very little per unit. However, when the standard constant high overhead figure was added the apparent cost was much higher than the prices quoted by competitive companies. The company's hot water heater listed prices were based on actual costs plus this high overhead so that their selling prices were too high to compete successfully with other outside concerns. Consequently the heater business was extremely small. A study of actual overhead costs indicated that the heater shop should use a much lower figure, not the average for the company as a whole. When costs were set on this basis, the unit heater selling price was reduced even below the values given by competitors. A lower price was quoted to customers and the heater business started to boom. This simple change in accounting methods and costing made it possible for this heater shop to quote prices that were reasonable, increase sales and show a profit for the company. Also the excessive overhead charges were charged against the proper jobs so that additional monies came in from research projects. This example emphasizes the point that cost data as well as quality information must be valid in order that the Reports to Management will have meaning.

Reports to and from Management covering items that affect the Quality Engineering group clarify the problems that must be solved. Through shorter and clearer reports better cooperation is obtained between all groups. If the lines of communication are simplified, maintained open at all times, then Quality Engineering is in a position to do a better job for the Company.

3. MISSING ELEMENTS

Management often seems so remote from the daily routines that those working on Reports to Management often fail to recognize items that would prove of the utmost interest and value to Management. These Missing Elements are woefully needed but are often not available. For example, in gathering scrap information a careful check may be maintained

in all areas to determine the number of pieces that are scrapped. Records are of such a nature that the tabulated items include only information by Station Inspection Operations so that a unit that must pass five Inspection Stations before going to stock or to the line is recorded five times. The actual number of completed pieces made in the Shop is not known, only the total number of units offered for acceptance at the various Inspection Stations. If the sum of these units is obtained it represents the number of Inspection Operations. If the sum of these Operations is used for a divisor for the number of pieces scrapped, where the average number of Inspection Station Operations per unit is five (5), the per cent of units scrapped is one-fifth (1/5) or 20% of the value for the percentage scrapped based on units completed for all operations. More frequently for many small runs of different piece-parts, it is too expensive to maintain complete records for so many different operations. The data should also contain a tabulation of the number of pieces made and sent to stock or elsewhere. With this information the percentage of pieces may be obtained. If this information is not available, this missing element should be supplied by a final check at a clearing Inspection Station that covers all completed units. The best information concerning scrap would be that based on loss in dollars for scrap at each Inspection Station Operation. Loss in dollars means much more to Management since a large number of pieces might be scrapped during the earlier operations and the actual loss in dollars might be a very low percentage.

Reports received from areas of all kinds will contain a lot of unnecessary information that could easily be replaced by valuable missing elements. The same philosophy holds true for the Reports issued to the groups by Management. It is often good strategy to keep from announcing an innovation before it takes place. However, Management cannot see the need of indicating in detail the responsibilities of a new department, the nature of the schedule, a proposed move from one building to another and the groups affected, the number of units that are made for research purposes only that differ from actual production units. In many cases no Reports are given by Management. This may result in many idle rumors such as: "The Motor Division is to be sold"; "A new Manager is to be appointed"; "Several recent contracts have been cancelled". The list of such possible rumors is very large. For maximum efficiency in operations, all groups should be made cognizant of the company plans as soon as possible so that time is spent on productive effort rather than attempting to determine company policies and what events will soon transpire that may affect the jobs and welfare of the employees.

Quality information both ways is sadly neglected. Reports are not made and little information is known as to Management decisions, customer returns, causes of failures, nature of reports from the field, Engineering action, corrective measures taken, and the success of these measures. Quality suffers from these Missing Elements.

4. REMOVING THE RUBBISH

Reports are generally too long and involved. The facts are cluttered together so that it is difficult to obtain the key elements in the report. Summaries are not given, or, where included, do not properly cover the items. The busy Executive finds so little information in these reports that he eventually pays little attention to them. Many find that a single one-page report will be read, whereas a carefully prepared report presented in two or more pages is not even considered. It is often difficult to pick the wheat from the cheff. How does the engineer preparing a report know what Management actually wants. It is an art to write a brief report that contains the salient elements and thus gains the attention of Management.

What steps should be taken to remove the rubbish? First, why write a report? Second, what is the objective of the report? Third, is it necessary to write the report? Having answered these questions and decided that there is something vital to cover, an outline should be prepared and the elements should be classified as to their importance. This is like the Classification of Defects, so familiar to the Quality Control Engineer. Are there any Critical items? Are some Major? How many are Minor, and how many merely Incidental? This analysis will show that only a few items are left that are of Critical or Major importance. Next, the Quality Engineer should put himself in the position of the busy Executive. Will he find the items selected for the report critical or trivial? If the elements pass this test successfully then the report should be prepared. If no elements are left after this strict scrutiny, no report should be made. Possibly a phone call might clear all elements.

Engineering Reports often consist of a page listing the findings that are covered in summary form. This page is the Report proper. Attached will be given the details upon which the report is based, arranged in logical sequence in accordance with the simple, summarized form given in the basic report. Even then it may be well to chart many items, tabulate many others, and omit many unnecessary details. Long sentences that are involved often lead the reader on an entirely different train of thought. So the Report gives concisely the subject covered, the nature of the problem, the solution, the basis of the solution, and the action sought. This cleansing of reports will reduce the paper-work and also lead to clearer thinking and a much better presentation.

5. HOW OBTAIN GOOD DATA?

The Conclusions in Reports to and from Management are no better than the data upon which they rest. What is meant by Good Data? It is basic data that contain a minimum of errors and pertain to the subject under discussion. Some data may consist of actual physical measurements; others, of test results. What are the errors in such data? The devices used to obtain these readings should have been calibrated. Their

errors should be known. If these errors are large, even larger than the discrepancies under consideration, the Report may be merely a consideration of instrument errors rather than a study of the difference between methods, materials, machine operations, performance, etc. Individuals may make many errors. Repeat measurements will provide some measure of such errors.

Much data may be based upon personal judgment. It will then be necessary to obtain men whose training is such that they will give valid data. Should you ask an electronic engineer about hydraulic measurements? Should you ask a dentist about bodily ailments known only to those in the medical profession? In Quality Engineering the judgment of individual inspectors is often used. What is Good Workmanship? When is a solder connection a poor connection? It has been demonstrated by experts in soldering that they can make a beautiful soldered connection without making the connection desired. Hence techniques for properly evaluating situations, requirements, machine parts, assemblies, soldering, etc. must be obtained. In order to properly do this, it is necessary to make use of all the knowledge available, and often necessary to call upon outside experts.

Good data will depend upon the standards set for obtaining such data. In the Shop, models, machine standards, examples of good workmanship such as welding and soldering, etc. will be set up and used as a basis for evaluation. The same is true in Assemblies. In market forecasting, a proper control must be established. Hence, one of the chief functions in Quality Engineering is to obtain the best possible economical standards for direct use. In X-Ray work it has been found that it may be necessary to establish two standards. One is a standard for taking the X-Ray covering exposure time and depth. The second is a standard that represents the poorest X-Ray reading that is considered acceptable. To obtain the latter, it may be necessary to check the performance of many border-line cases to determine whether they are satisfactory for use and thus establish the limit indicating the beginning of sub-standard quality.

Planned experiments will have to be developed for obtaining the data in the best possible form. Sample results are used as representative of the phenomenon under consideration. Care must be taken to obtain truly random samples from the total population of events, parts, etc. being studied. If data can be obtained that contain minimum errors, it should be possible to repeat the procedures followed in obtaining these data and secure the same results within a desired narrow band. Checks for significant differences may be conducted.

These data may then be summarized and filed so that they may be referred to readily. Such data then are the basis for the Reports that are to be made. These original data, together with a listing of the conditions under which they were taken, should be filed. They should be maintained for ready access when points in the Reports to Management are questioned.

Unless the data are controlled within a reasonable band, no valid predictions may be made. Hence good data are a necessity for Management Reports.

6. FILLING THE GAP

Management and many in lower echelons will not believe there exists a gap in their administrative structures. Most Managements believe their system works and has no Missing Elements. The Quality Engineer, however, soon finds a lot of missing items. Data sheets are only half-completed. Summary Reports are made only upon request. When made, they often cannot be completed due to the lack of sufficient good data. A month may pass without any action being taken to relieve this missing elements condition. The Quality Engineer must bridge this gap. This group must set up procedures to provide inter-related data that have meaning. It must be in such form that it can be computed readily. There is no necessity in taking too much data but there is sense in using standard forms that are easy to follow. There should be a common pattern from form to form so that there will be a minimum of error in these data.

Management wants reports backed by facts. Data must be in such form that they can be explained and also will substantiate the reports. When the reports and the data seem to differ, the validity of the reports may be questioned. Hence Quality Engineering must set up stream-lined forms with simple directions for their use, that give all the essential information, also sufficient supplementary data that may be used in special Management Reports. The time required for making out such data sheets may be greatly reduced if numerical cells are provided for tabulating measurements by checks rather than requiring each measurement to be written. Every part should have a meaning and every line should be completed in most instances. The inspector or tester will then be certain that he has completed the form. If there are many lines to leave vacant except for special circumstances then there will be many gaps that will be discovered too late. Missing data will be avoided by the use of scientific forms that are well designed.

7. HOW PRESENT DATA

There is an art in presenting data in such a manner that the reader immediately becomes interested in the picture and quickly gains the concept that is being presented. Sometimes the written word is used and sometimes charts and tables are added. Many excellent references might be given discussing good charting practices. Such texts should be studied and also charts and tables prepared by the various groups in the Department of Commerce (such as the Bureau of Standards), Research groups in Banks, Travel Agencies, Technical Tax groups, and the Statistical and Operation Research Departments in the large industries. Specialists in these fields have clever devices for attracting attention to those points that are emphasized. Advertising agencies present attractive

brochures that attract attention and have appeal. They are constantly making improvements in their presentations. The engineer too often feels that a simple statement of the facts should convince the reader. This is not true. It requires salesmanship of the highest order to prove conclusively to Management that certain actions must be taken immediately.

One engineer stated that he had completed a study whereby it was obvious that the company could readily save $100,000 annually by making a few simple changes. After presenting his case he was much chagrined to find that no one seemed interested in his work and these apparent savings. There must be a dynamic approach to the problem. The facts are repeated in so many ways that the busy executive finds it impossible to escape considering the case.

One of the basic rules is that each chart should be self-contained. All keys to terms used should be on the chart. Also its line structure should be pleasing and arresting. The center of interest must be evident almost at a glance. Tables also should not require text material to indicate their meaning.

The presentation must be in simple form. It should be as short as possible. It should preferably emphasize one point only. It should be in such form that, if action is required, the nature of the action is clear. Some reports are merely informative. One form of presentation that has been found extremely valuable is to tabulate the data and note such brief reading material as is needed on one sheet and on the opposite page on a second sheet, visible at the same time, present a graphical chart of the data. Some executives like visual presentation; others like tables. This dual presentation satisfies both.

8. SCHEDULING REPORTS

Management will allow some lag in the presentation of certain information. In other cases the data must arrive within a few hours after it has been obtained in order to be used to the best advantage. If you are speculating in the stock market, information that a given stock will soon declare an extra dividend means that you should purchase this stock before the information is common knowledge; otherwise the price will increase so much that no profit may be obtained from this early tip. Minutes may make a difference between obtaining a large profit or practically none.

On the machine shop floor it is necessary to feedback information rapidly if it is found that a machine making parts at the rate of 50 per minute is making a part to the wrong dimension. The sooner that corrective action is taken the less scrap is generated.

A careful schedule should be worked out with the different branches of Management so that Reports will arrive when they will do the most good. Sufficient lead time must be ob-

tained to do the necessary analysis, tabulation and typing. Where the lead time is extremely short, format should be prepared in advance so that the data may be inserted rapidly. A good rule is to keep charts as current as possible. Also make the forms used for recording in such a way that carbons of original data may be used directly, wherever possible. Extra transcribings of data require more time and introduce errors. Machine and IBM runs should often be used to speed up tabulations and computations.

Make a reasonable and definite schedule and adhere to this Schedule.

9. WHAT AN ADMINISTRATOR SHOULD KNOW

This topic may seem impertinent but is given for contrast purposes. Administrators should learn from their Quality Engineering department what information is available under current conditions. They should know what additional information may be obtained at a given cost. They should know what information is available in other departments to see if there is an overlapping of data, or worse yet, a different procedure followed in obtaining certain types of information so that the final results will be contradictory.

The Administrator should know what information he needs at certain times and the effort, lead time, and cost required. He should reduce his demands to something reasonable in line with budget demands. All possible short-cuts and efficient means for handling data should be permitted to be used even though several different departments must work jointly on the desired Reports. A program should be established in line with such studies and conferences between all groups so that the task for the Quality Engineer is well defined. Also the Administrator should indicate what Reports issued by Management will be made available and when. The good Administrator works with and through the Quality Engineer. The result will be a strong and well co-ordinated program.

10. WHAT A QUALITY ENGINEER SHOULD KNOW

Quality Engineering requires the utmost diplomacy. It is necessary to know how all the other groups function within their own departments so the Quality Engineer can contribute most to those areas where they must, of necessity, consult with him and obtain his decisions. The Quality Engineer must know in general the technical details related to all products made by the company. Where these vary greatly, sometimes only products from specific shops need to be known. He must know materials, processes, methods, statistics, physics, production, inspection, mathematics, specifications and often Military requirements and documents, contracts, inspection and test methods. In particular, he must be a Specialist in an area he likes best. Above all he must be quality-minded.

It is absolutely necessary that the Quality Engineer know how to prepare meaningful Reports. He must know how to

analyse and summarize data in such a way that Management will fully understand the text and context. He should know how to prepare simple forms, yet complete enough to provide all the data required. He should know how to make clear, self-contained charts. Also he should know Design of Experiments, Analysis of Variance, Correlation, some aspects of Operation Research, so that he will be able to assist Research and Development, Production Engineers, Process Engineers, etc., and also provide aid to Management in the field of Management. He should be able to assist Management in formulating problems in such a way that solutions may be obtained through the use of Operation Research techniques. He needs a sixth sense to secure the factual information and comparison data necessary for Management to make the best possible decisions.

In addition to these qualities and talents, the Quality Engineer must be a good supervisor. He must be willing to take orders, and in turn must be able to outline clearly to subordinates what is required. He must be able to work with people in the shop and also those in Management. It may not be possible to obtain all these excellent traits and abilities in one individual. However, a group of Quality Engineers may be put together, similar to an Operation Research Team, that will have as a composite most of these abilities. With such personnel it will be possible to establish Quality Engineering on a true Professional basis. Management will have a high regard for such a group and will give them strong support as long as they keep within their field. A rapid survey of such a group is obtained by checking the Reports they issue. If these are simple, clear and concise, and also are in demand by other departments, then it may be felt that the Quality Engineering department or division is successful.

11. OVERLAPPING FUNCTIONS---RESPONSIBILITIES WITHOUT AUTHORITY

From the description of Administrative and Quality Engineering functions in the two preceding sections it is clear that functions in any company assigned to any one group must be carefully differentiated and separated from the functions in other groups. There is a tendency in many companies to have a number of groups that work along very similar lines. Their duties tend to overlap so that some action started by one group may be partially covered by another group. Engineering may have a Reliability group. At first they just assist the design engineer, but after a short period of time they take over policing of production lines "to obtain good engineering information" of course. Their actions may often be in conflict with the independent Quality Control group. There may be a Statistical group or an Industrial Engineering group that finds itself under-loaded so includes without consultation or assignment reports on the activities of other departments. After a time two or three groups have many functions and work in process that are essentially alike. To overcome this it is necessary to clearly define the work of each group and the limits of their functions and also their activities.

When the duties of Quality Engineering are not clearly defined by Management, the stronger groups with somewhat more authority will take over some minor Quality responsibility. Where Quality Engineering has the task but no authority to perform the task, it will be glad at first for such cooperation. It, in turn, may be assigned work that should have been given to another department. Engineering is supposed to run the Qualification Tests and provide Quality Engineering with the data. Also it is advisable to assign a Quality Engineer to observe all qualification tests. Who then must prepare the test data and test procedures in accordance with the requirements of the customer? The Engineers may expect Quality Control to do this, or may elect to do it without checking with Quality Control. Since these Qualification Tests are so nearly like the tests in production and inspection and also must be performed by Quality Control on a periodic basis, then Quality Engineering should evaluate all such tests. It should also work with Engineering on standard forms for use in preparing such data for the use of customers. Customers, especially the Military, look to Quality Control groups for quality information on models and prototypes as well as final production units. If Quality Control is held responsible for the final acceptance of units, together with Inspection, it must have the authority to obtain all possible test data on the project and also be in a position to conduct independent tests for verification.

Overlapping functions must be avoided. Authority must be assigned to the one given the Responsibility for performing the function. Quality Engineering must carry its full load. When these functions are distributed to other groups, many are never covered, contracts may not be met, Management suffers, and Quality Engineering loses its Professional status.

12. MAKING A PROGRAM TICK

If the lines of communication in an army fail, the strategy planned will fail. The same is true for a Quality Control Program. The System must be simple but effective. Lines of communication must be clearly outlined. Forms for obtaining action for all the functions must be complete and provide all required information in a manner that makes it possible to use in later reports. The liaison with other groups must be maintained but must operate through designated channels. All Reports from Management must be sent directly to the Quality Engineer responsible for the action requested or who is to use it in the preparation of a composite report later.

The Reports to Management must be so prepared that they indicate the progress that is being made and the decisions reached. They should tie the program together so that any outside concern that wishes to evaluate the program will know that effective controls are being maintained on all Quality functions. Customers must feel that the quality of their purchased products is being monitored at all times. Rating systems should indicate where controls are strong and where, weak. When there is true teamwork between all groups and all areas are operating smoothly, then one may be certain that the

Program ticks and is truly effective. Cost analyses will corroborate this. Such checks will show that quality is being maintained at economical levels with minimum expense.

13. ULTIMATE GOALS

If every function was performed without error, the designs were perfect, production made all parts to print, systems were assembled perfectly, all tests were met with wide margins, then the outgoing products should approach 100% performance. If studies indicated that test, performance, field, shelf and storage life met all the demands of service, then there should be 100% Assurance that the Reliability was 100%. This ideal condition is the ultimate goal.

Since perfection cannot be attained, Quality Engineering may be satisfied if a large percentage of the functions are under control so that inspection and test may be minimized. The program will provide valid measures of Reliability for components, sub-assemblies, assemblies, and the System as a whole. A minimum amount of testing will be required and there will still be maintained adequate Assurance that the Reliability specified is more than met.

14. GETTING INTO THE ACT---MAKING DECISIONS

The last two sections have presented an Utopia. Since Quality Engineering is far from these goals and many programs are in trouble continually, it is necessary to see what can be done by Quality Engineering. Management must be made aware of the important part that Quality Engineering can play in a good Management program. If good Reports present all phases of the activities of a strong Quality Engineering group, all groups will be forced to arrange their programs so that they may assist in the effort to reach the ultimate goals. If the Quality Control group is weak and is doing nothing but trouble-shooting, then the Reports should be made in such a manner as to sell a true Quality Engineering program to Management. This program should include all functions normally included at other companies under Quality Engineering in order that savings in inspection and test costs may be realized and that finally the Quality Engineering Division is firmly entrenched in the organization of the Company.

When Quality Engineers are found to be sufficiently mature to be completely trusted with their responsibilities, then they should assume more responsibilities step by step, and check the validity of the decisions they make. In the past Quality Control and Inspection have been considered merely as service functions. If true Professional status is to be obtained, then they must accept certain responsibilities. The executive that is most successful is the one who makes the least errors in his decisions. He is an executive because he is willing and able to make decisions. Quality Engineers are usually not in that position. Too many times Quality Engineers are permitted only to recommend while the decisions to be made are left to other groups, such as Engineering, Production, and Purchasing.

When the day comes when Management recognizes that Quality Engineering has an area for which they are responsible and that it is capable of making decisions the same as other groups, then there is beginning a little spark of recognition for the group. Reports to Management must reflect this so that Quality Engineering may quickly develop into a Profession and not be merely a service organization.

15. MEASURING RESULTS

Quality Rating Charts were mentioned earlier. Reports should include measures of performance for all the departments and divisions. Numerical measures of performance should be obtained. In all inspection areas some form of Classification of Defects should be used. Some will be simple; others will be more complex. The Demerit Rating System is one of the best. It may be used for both Attributes and Variables data. More variables data must be obtained and more uses made of such data. Measures of Reliability with given Degrees of Belief or Assurance must be obtained. These may be achieved through the new Positive Approach to the Reliability problem, wherein by almost continuous analyses a pattern of Reliability is obtained much earlier and at minimum costs.

Quality Reports may be condensed into Rating Charts that will present true measures of performance. All functions must be covered. The truly effective Department will be the one that shows Management in Dollars and Cents the savings they have effected, shows the cost of their operation, shows what can be done to obtain more savings and presents a program for future growth in line with past performance. Obtaining numerical values for the intangible items will require much ingenuity. It can be done. Management would like their Reports to reflect Results attained.

16. MAXIMUM BENEFITS FROM MANAGEMENT REPORTS

Consideration of the factors covered in the preceding 15 sections shows that Quality Engineering has a long way to go before it can show Management that it truly has Professional stature. The goals are there. The attainment of these goals has been slow and sometimes negative.

Maximum benefits will be obtained by Management from Reports submitted to them when they are placed in constant use. If these Reports are considered valuable, then, when a Report is not received it will be missed and a request will be received for the missing Report. Each Company will have different types of reports to fit the needs of each. Those which will be demanded by Management will be the Reports that provide a bird's eye view of all activities in the Company with very little effort required from the executive. Other Reports providing factual information from which Management may reach proper decisions should be prepared so that they will be used as one of the key factors in reaching final decisions.

Many companies have excellent Reports prepared regularly by their Quality Engineering group. They portray Customer Complaints, and Current Quality for all products from all shops, departments and divisions. They indicate completed action items, progress of those now under consideration, and contemplated actions. They may even show market results for competitors and some measure of their quality in comparison with the company's own products. Use is made of Summaries, Tables and Charts. In one company a weekly book of standard reports was condensed to less than ten (10) pages. Management must see its entire field of activity quickly and clearly hour by hour and day by day. However, weekly, monthly, quarterly and yearly results must be used as adjuncts to the daily reports. They make a successful reporting system complete. Trends and long range planning are based on these Reports. The Reports, above all, must not be too elaborate and must be presented in concise form. Their form should be such that they may be used in making forecasts of future performance.

Reports that are used and in demand will be those that have Maximum benefit to the Company. Few are trained to prepare such gems---factual, clear, accurate and concise. Each Quality Engineering group should strive to provide the best compact Reports possible for their Management. If they are used, such efforts are truly fruitful. Quality Engineering then may use Management Reports as a means of obtaining true Professional Recognition. If such standing is attained, then their Reports will in turn for many features be truly Reports from Management.

REFERENCES

1. Flesch, Rudolf, "The Art of Plain Talk," Harper & Brothers, New York, 1946.

2. Flesch, Rudolf, "The Art of Clear Thinking," Harper & Brothers, New York, 1951.

3. Kent, John L., "Engineers Can Write Better Technical Reports," Chemical and Engineering News, Vol. 33, p. 444, January 31, 1955.

4. Kent, John L., "Improving Technical Communication Techniques," Education-for-Technical Communication Conference, University of California, Los Angeles, December 1, 1956.

5. Spear, Mary Eleanor, "Charting Statistics," McGraw-Hill Book Company, New York, 1952.

STATISTICAL AIDS TO MANAGEMENT

Morris H. Hansen and William N. Hurwitz
Bureau of the Census

One of the major functions of an executive is that of making decisions. Although a large proportion of these decisions can be based on information, many of them are based on intuition. Given sufficient factual information, many of the decisions that otherwise are time-consuming become automatic and need not divert him from issues in which qualitative judgment is essential.

Most decisions which concern the executive can be reached more quickly and effectively by a knowledge of relevant facts. For such decisions the necessary information may concern the internal operations of the business, the activities or characteristics of his customers or dealers or of his potential market, or even more general external types of information. The required information may involve experimental work for evaluating the effect of alternative methods, selections or treatments. It may involve operations research, and require formulation of models and the evaluation of expectations and probabilities associated with alternative events. Sometimes the problem is one of assembling statistics from existing sources; often it involves the collection of completely new information.

In assembling relevant information or evidence the statistician has methods at his disposal that can be applied with great effectiveness in acquiring sufficiently accurate information at low cost and with great rapidity. These methods include, among others, modern sampling, principles of experimental design, and quality control. Statisticians also contribute to executive decisions by examining problems in collaboration with subject matter specialists or administrators to determine the kinds of factual evidence that may be of assistance, and formulate data collection and compilation methods to yield needed information. In any of these situations, the statistician carries the principal role in the design of methods for assembling information and guiding in its interpretation.

As an example, the data on the internal operations of a modern business or governmental organization are usually so well recorded that there are tons of paper reflecting its activity. In many instances, much valuable information is buried in these tons of paper; but synthesizing the information into some useable form often becomes a matter of great cost because of the sheer magnitude of handling this volume of paper, containing thousands or even millions of numbers. The individual pieces of paper may have to be classified into meaningful categories - arithmetical operations may have to be performed on them; or they may have to be reviewed for completeness and consistency. On the other hand, more difficult problems are often presented by the many activities of an organization on which records or descriptions are not available, despite the tons of paper work that are created. Direct observation or collection of information must be instituted if factual information is needed. Perhaps these remarks are sufficient to indicate that the process of obtaining information within a business organization is in some ways similar to the chief function of the Census Bureau.

Since our major produce is information, we have had considerable experience in collecting and processing data and have given a great deal of attention to applying and creating better methods in this field. We consider one of the most important steps has been in the direction of achieving an atmosphere in which creative thinking can thrive. We believe that this has been facilitated through the form of organization structure we have evolved at the Bureau.

An organization in existence over a period of years tends to develop traditional ways of performing its functions. Improvement is often slow and evolutionary. However, major steps forward must of necessity break with tradition. The classical organization structures in which research personnel have staff status are not always very effective in utilizing the results of research because of natural resistance to change on the part of operation personnel. On the other hand, it is very difficult to imagine research in a line role. Without the interchange of ideas between operating and research staff, however, traditional methods often become embalmed and progress stymied.

In terms of our own experience, we have evolved a hybrid organizational structure at the Census Bureau which represents a break with classical traditions of organization theory. We have a Statistical Research Division which has primarily a staff function and almost every one of the subject-matter divisions has a Statistical Methods Branch with a line function. The personnel, although principally mathematical statisticians, includes psychologists and other specialists. These people are not interested in theory alone, but in the application of theory to specific operational problems. The most important characteristic of these specialists is a versatility which not only gives them an interest in a large variety of problems but also permits them to deal competently with most of them.

The distinction between staff and line functions of the Statistical Research Division and of the various Statistical Methods Branches is not a rigid one. Moreover, each Branch chief has a dual responsibility; he is administratively responsible to the Chief of his subject-matter division and technically responsible to the Chief of the Statistical Research Division. Although this may seem to violate principles that have been enunciated in classical texts on organization, modern theories of organization have been gravitating toward acceptance of this concept. In our own case, we have found that this kind of system works extremely well.

In many respects not only the research group but the entire staff of the Bureau of the Census works on our research problems. This, in a large measure, is a direct result of the organizational structure of our research function. This is true because of the give and take between those responsible for programs and operations and the research group who tend to be the exponents of change. Since members of our research team do not have a day to day operating responsibility they are free to ask questions directed at improvement of programs and operations. Theirs is a responsibility to ask questions. Of course they have the authority to do so as well; but I emphasize responsibility to ask questions. The important thing is that the research group is a versatile group with a duty to inquire into what is behind what we are doing, to learn why we are doing it, and to see what techniques are available or can be developed for a particular job.

Most of us are aware of the confusion which might arise if a group with primarily a research function were permitted to communicate its findings directly to the production line in the form of orders. In our judgment, the things that make this approach exceedingly effective are: (a) the central research group cannot issue orders but works through persuasion; (b) the high caliber of the research group; (c) the free exercise of the authority to ask questions without fear of appearing to be naive in raising foolish questions about things that have long been "proved" by custom or tradition, (d) the recognition that criticism is most effective when accompanied by constructive recommendations or alternatives; (e) the high caliber of the executive staff of the Bureau and its strong support of advanced methods and good administration; and (f) the presence as a member of the executive staff of the director of the research group. While the executive staff of the Bureau of the Census are strong supporters of advanced methods, the burden of proof that a new proposal is superior tends to be on the personnel advocating the change. The principles we have outlined provide the necessary stability for proper functioning of our organization, but with a sensitive reaction to potential advances that may be available.

As a result of the interaction between the research group on the one hand and the program and operations groups on the other, the statistical research program has had an important effect on Census methods, not only directly but also indirectly by stimulation of other groups within the Bureau. As a result of some of this research Censuses and surveys are now made in ways which are significantly better than those used in the recent past. Many of the methods we have developed or applied are potentially applicable in areas where statistics can be of help to management.

Let me illustrate by describing some of the applications of statistical methods to Census problems developed jointly by the research groups and other members of the Census staff. In developing and applying these methods our purpose has been to achieve a product that is of the desired quality, at minimum cost. In some cases, the proper balance of the expenditure of resources with the quality of the product means that the accuracy should be decreased at some points, and increased at others. We feel that it is not wise to turn out information which is more precise than is needed. It is frequently true that important savings can be made because the marginal gains in accuracy that can be achieved after certain levels have been reached often come at very high prices.

Among the activities of the statistical research group of the Bureau, sampling research and applications occupy an important position. Sampling is the field in which this research organization began its work, and from which much of the basic philosophy of our program arose. It was the extension of the philosophy of statistical work that was created in sampling research and applications that provided the basis for the present state of development of the design of efficient surveys and of the measurement and control of errors in surveys. In sampling, one of our basic principles has been to use methods which yield objective measures of precision of the results. This is accomplished through the use of probability sampling methods. By using these methods of sampling we can obtain from the sample itself a measure of the magnitude of the difference between the results of samples of the size and type used, and the results of a complete census taken under essentially the same conditions. When we limit ourselves to probability sampling procedures, the

mathematical theory is a powerful guide in the effort to achieve the maximum precision of results per unit of cost.

With the aid of this versatile tool, more information can be obtained for a given cost. As an example, in our Decennial Population Census, basic data like sex, age, and race, are collected from all people in the United States. Information on income is collected only from a sample of the population. Sampling permits us to obtain the information at a reasonable cost and on a faster time schedule.

Sampling is applied as an administrative control in the Bureau as well as for the collection of sample survey or Census data for publication. A sample can furnish more timely information than can be obtained from the complete processing of records. Running expenses or the nature of operations of a department can be obtained almost up to the minute by a sampling procedure whereas accounting processing of all the records might be too expensive and badly out of date. As an example, during our 1948 Census of Business we had 308 field offices supervising some 35,000 enumerators collecting information from business establishments throughout the country. We had a procedure for collection which represented an improvement over prior procedures. There were still some questionable aspects of the final procedure, however, which required that we keep a close watch on the operations so that last minute changes could be made. In order to do this two small samples of offices were designated. Expenditures in a sample of 15 offices were watched through sample time studies of enumerators and office operations. At the same time the procedures were critically observed in another sample of 20 offices. On the basis of very early returns from these samples, changes were made in the procedure which enabled us to complete the job within the budget allotted. Had we not had timely information from these two samples, we would have learned too late that we were spending much more money than allotted for field operations. This, of course, would have resulted in a curtailed publication program of business statistics. We have found numerous opportunities for such effective application of sampling.

A paradoxical advantage of sampling is in the improvement of the final accuracy of measurements. It is well known that sampling errors are introduced into estimates compiled on the basis of sample returns. It is not so well known, however, that under certain circumstances the basic data can have a higher degree of accuracy because it is possible through the use of savings due to sampling to concentrate on obtaining better accuracy on the basic data. An example of this effect is afforded by the 1948 Census of Business. This Census included a series of questions on commodities sold. It is difficult to get good responses to these questions because usually records do not exist in the individual retail establishment that give sales by commodity or by commodity class. In past Censuses of Business we found that we had to stop a good deal short of getting complete schedules when we tried to collect this information from every establishment. We found that after we had decided to ask these questions of a sample of

establishments we increased the response rate for these items from about 60 or 70 percent to more than 90 percent. Because we were dealing with a sample we were able to insist on callbacks and to take other steps to increase the accuracy. Thus we were able to get considerably more accurate results.

An area which is of interest where basic information is subject to the vagaries of human memory, estimate, or conjecture is that of response errors. The measurement and control of response errors in censuses and surveys is one of our newest developments. In particular, one of the most note-worthy activities is the effort since 1945 to evaluate the accuracy of the results of the major censuses by means of sample surveys. Starting with the 1945 Census of Agriculture we have adopted the policy of making such a post-enumeration survey for every major census. This work involves the selection of an appropriate sample and a revisit of the sample households or establishments by carefully selected and highly trained enumerators. These enumerators are paied by methods that permit them to spend more time in carrying out their work and to use extremely intensive questioning on selected topics, in order that we may be able to evaluate the accuracy of the data originally obtained.

There are two purposes in carrying out such a post-enumeration survey. One is to determine the errors that occur and their sources in order to develop economical corrective measures in future surveys and censuses. The other is to evaluate the magnitudes of the errors that occur in order to guide the users of Census data. On the basis of knowledge of the reliability of the results, users should be able to employ these data in a valid way without attempting to draw inferences or to formulate policies not justified by the level of accuracy of the basic data.

The surveys that yield the measures of the quality of the Census are not themselves perfect, just as the Census is not perfect. However, we can and do employ in the post-enumeration surveys much more accurate methods than is economically feasible to use in a complete census. We have also made it a practice to evaluate the quality of the post-enumeration survey itself by means of a subsample that is inspected by members of our own staff, and by means of a careful examination of various aspects of the survey by the staff.

Another way in which statistical methods can be applied is in the control of the quality of information processing operations. Traditionally, the office operations on the schedules collected in surveys and censuses had been carried out with the notion that perfection must be attained in each operation. This has usually implied the complete verification of every operation, and sometimes more than a single verification of an operation. Our philosophy in this connection now is that we must have a rational balance between costs on the one hand and accuracy on the other. By the introduction of verification on a sample basis and the use of statistical quality control methods, some standard and others new, we have attempted to achieve this balance. This has meant decreasing the accuracy of the office processing to a level commensurate with the accuracy attainable in other aspects of processing. The methods of quality control that are used are such

that they give results of specified accuracy but at considerably lower cost and with a faster time schedule than would be possible otherwise. As a consequence, we have been able to transfer some of the savings into the work of improving quality where such improvement is needed. The principal area in which such improvement is needed is in the original collection of the data.

A sampling application which has been known for most of this century but only recently has been getting the attention it deserves is known as work sampling or ratio delay. By sampling in time, it is possible to obtain information on operations quickly and effectively which would be difficult or impossible to get without sampling. One of the most useful applications of this method is to obtain a distribution of the time required for various elements of a job. This information frequently leads to redesign of an operation with the proper stress on its essential aspects.

A tool with which many of you are probably very familiar is design of experiments. It is possible to evaluate through experiment alternative procedures, methods, and equipment. In spite of it being so well known, it is worth mentioning that it can be used in the field of information handling as well as in manufacturing operations, although on occasion comparison of the results of human effort may be quite difficult or impossible to interpret.

A method borrowed from the physical sciences and becoming more popular today is the construction of mathematical models. The mathematical model represents a process by characterizing its essential features in mathematical symbolism. Its great advantage lies in its synthesis of the characteristics of the process into a formula or set of formulas which the mind can encompass and understand. In this manner, it is possible to study how the process can be controlled or manipulated. We have done some work in this area in creating models of verification processes which enable us to understand the process and to make better decisions regarding methods of quality control. We have also constructed a model of Univac operations which enables us to decide more economically when to consider the equipment out of service and to evaluate and control the effectiveness of utilization.

I am sure that most of you are aware of the mechanical devices available for processing information. One of the areas of research in which we have been engaged is the development of mechanical methods of tabulation, computation, and control of quality. Historically, the Bureau of the Census participated in the initial development of punch card tabulating equipment. Hollerith and Powers did their early work at the Bureau of the Census. We have our own mechanical laboratory and machine shop where we develop, construct, and maintain some of our tabulating equipment. We use commercially available equipment as well. A comparatively recent development with which most of you are probably familiar, is the development of large scale, high speed, electronic computing equipment. The Census Bureau participated in the developmental work in the application of such equipment to large scale data processing, and paid for the design and construction of the first Univac. We now have two Univacs in service. The availability of such equipment means reduction of costs, acceleration of work, and increased accuracy of results.

Modern electronic computing equipment is exceedingly versatile, and displaces some operations formerly done manually as well as those formerly done on punched card equipment. One such use is to displace a form of quality control called critical review or "editing" of returns or tabulations. During the 1950 and 1954 Censuses, increased use was made of such methods, thus eliminating some manual processes. For example, it has been customary to subject schedules to editing processes for internal consistency, making certain adjustments when they are not consistent. During the past few years we have been doing some of this work mechanically, leaving inconsistencies on the punch card but using the mechanical tabulating equipment to identify the inconsistencies and to permit disposing of them in accordance with specified rules. Such methods have provided a way of decreasing cost and improving the timeliness of the final results.

Having had some experience with the versatility of modern electronic computers, we are moving more and more in the direction of quality control on the computers. More editing work can be done on electronic equipment than has been done heretofore on mechanical equipment or manually. Moreover, most recently we have been moving toward the elimination of manual verification of coding and substituting a coding verification by computers. This can be done only when it is economical to feed to the computer some or all of the information used by coding clerks. Although the computers do not catch errors with certainty when the verification is done on the basis of incomplete information, we believe that the error rate of the new system will compare favorably with that of the old one for a given cost. Conversely, computers make no clerical errors to which manual operators are highly subject. We expect many future advances in the application of computers to control of remaining clerical operations.

The examples of the use of statistical methods we have enumerated do not begin to exhaust the potential. Because our time is limited, we cannot cover the entire field. In our opinion, however, the Bureau's largest single contribution has been the creation of an atmosphere which makes highly competent statisticians willing to forego the more obvious advantages of higher income outside the government for the more subtle advantages in our organization. These advantages are the wide range of challenging problems requiring imaginative thinking; the opportunity of a serious hearing no matter how extreme the idea; and the feeling of having made a worthwhile contribution when a brainchild is put to useful work. As long as this is possible, we foresee our approach continuing to generate in the years to come a true science of the collection and compilation of statistics in general which will be potentially of great help in aiding management in its decision making process. Moreover, the organizational approach we have taken is, we believe, generally applicable and can aid in providing effective statistical aids to management in diverse activities and organizations.

ARMY QUALITY CONTROL AND INSPECTION

Brig. Gen. A. F. Cassevant
Department of the Army

Introduction

I welcome the chance to speak this morning on a subject that has troubled the Army for some time. Too often people are content with old programs that appear to accomplish their mission and give little thought to the question of whether or not the programs are keeping abreast of changing conditions. Abe Lincoln once said: "If we could first know where we are and whither we are tending, we could better judge what to do and how to do it - -." In our most recent attempt to keep ourselves abreast of the changes occuring around us everyday, we have reviewed the Army Quality Control and Inspection program from three viewpoints: first, why are we doing as we are; second, what does Industry - the other member of the Army-Industry team - think of our program; and third, are we operating in the most economic manner possible that will do the required job. In each of these three areas our review uncovered needed changes. Before telling you about the Army's new look in Quality Control, a brief review of history and current concepts is in order.

History

Inspection and Quality Control have progressed immeasurably since the days prior to World War II. At that time, most inspection programs were considered necessary evils. In a few places however, experiments were being made with sampling systems, examples being the full fledged quality control and sampling inspection activities conducted at Bell Laboratory and Western Electric. Today, in the military departments and in many places in industry, regular use is made of quality control techniques unheard of in the 1930's. Undoubtedly, most of you gentlemen are aware of these facts and probably have discussed the topic in detail many times. One aspect of the change brought about by placing emphasis on quality control and inspection leads us to briefly discuss, what I consider, an important advancement. This is our ability to better understand what specification and contract requirements actually mean and what they can do for us.

Man, all through history, has disliked changing his ideas and beliefs. This has been particularly true in his belief in absolutes. However, he has had to change. For example, not too many decades ago man was positive that he could order the making of something, a steel pin maybe, exactly two inches long. Of course implicit in this belief was also a sureness that things could be exactly alike. Such beliefs were not unreasonable, when you consider the fact that no one could measure precisely enough to distinguish the differences between items believed to be exactly alike. Man's ego led to a desire to believe in perfection, especially perfection created by man.

We have come to accept the impossibility of absolutes in terms of size, hardness and similar properties of things. For the most part, we seem reluctant to come to grips with the problems of mass of details and complexity. Our inspection systems of twenty years ago did not measure quality of product in the aggregate with precision. Today, Quality Control Techniques measure such quality with high precision. These

techniques lead us to understand that seldom is there perfect quality in terms of absolute compliance with specifications, drawings, and other contract requirements. I suspect that here is another absolute, falling-by-the-wayside, and especially will this be true, if we require economy and efficiency. To require perfection in the sense in which I now use the term, when we know from experience that we don't need such perfection, amounts to pure wastefulness. During World War II we bought, and used satisfactorily, billions of dollars worth of supplies that ranged from 1/4 per cent to several per cent defective in terms of absolute compliance with specifications, drawings, and other contract requirements.

The precision of our new tool, "Quality Control," has in one way created problems. As with any really sharp and unfamiliar tool, people sometimes get cut while learning how best to use it.

During the days of inspecting by spot check, and accepting or continuing inspection to sort out rejects, the need for explicit contract statements regarding the inspection procedures to be used was not apparent to most people. A few people were ahead of the game however. For instance, about twenty-five years ago Mr. Bancker, Vice President of Western Electric Company, defined specifications during an American Society for Testing Materials meeting as follows: "Specifications are simply definitions of the properties which the purchaser desires. In them are his best efforts to state, in measurable terms, properties necessary for satisfactory use. They include test methods . . . and provide an agreed upon basis of inspection." More recently the National Security Industrial Association in a report on Inspection and Quality Control stated: "It has been observed that most contracts are subject to changes and amendments, but very seldom are corresponding changes negotiated in Quality Control procedure or assurance requirements. They too should always be subject to review and amendment as conditions justify."

These examples, and others that have come to our attention, point to a growing awareness on the part of both industry, and the Army, that inspection and Quality Control procedures are now sharp tools, the use of which must be agreed upon in advance of purchase contracts. Without advance agreement, one or both of the parties concerned may suffer because of a misunderstanding.

These changes in the fundamentals of Inspection and Quality Control have led to experimentation on the part of industry and government agencies. Methods of purchasing are being revised, especially as regards vendors' quality control systems. I fear that in some instances the changes in purchasing practice have been further reaching than is necessary or desirable. There are pitfalls in racing to keep abreast of the times. Military purchasing activities are in danger of becoming responsible for contractors' internal management. This is due to the newness of quality control and to the fact that top management has perhaps not had time to become acquainted with its basic principles.

Functions of Army and Industry

Before discussing Army Quality Control policies and procedures in detail, I would like to dwell a little on the division of functions and

responsibilities in the Army-Industry team.

Army subscribes whole-heartedly to the concepts set forth in Department of Defense Instructions which read in part: "The basic quality assurance concept of the Department of Defense is predicated upon the fact that: (a) responsibility rests upon the contractors and producing activities for controlling product quality and for offering to the military departments for acceptance only those items or lots of items considered by them to conform to contractual requirements; and (b) responsibility rests upon the military departments for determining that contractual requirements have been complied with prior to the acceptance of the product."

This division of responsibilities, if fully implemented by both Army and Industry, is one which will lead to a most effective military supply system. Industry, in a free economy, is particularly well suited to decide the best ways and means of accomplishing the economic production of supplies, whether they are commercial types for competitive sales, or for use in waging warfare. The responsibility, of a military department making a procurement, to insure that contractual requirements have been complied with, does not in any way include responsibility for the management of a contractor's internal affairs. Army policies and procedures have been dictated and established with this concept in mind.

Army Policy and Procedure

As I have stated, Army policy dictates that in every way possible the vendor should remain autonomous in the management of his business. This should in no way reduce the nation's strength during mobilization, since our vendors' day to day commercial pursuits require them to maximize return for dollars spent. Dollars of course represent nothing more than man-hours and materials, the very things that we must conserve during a period of mobilization. For this reason, the Army Quality Control policy governing items which can be factually described and for which we can write definitive Quality Assurance provisions considers that vendors are responsible and have full authority. When such specifications are possible, any other course of action violates a basic principle of all organized human endeavor. This principle is well stated in the "Dictionary of United States Military Terms for Joint Usage." In this Dictionary the word "responsibility" is defined as follows: "The obligation to carry forward an assigned task to a successful conclusion. With responsibility goes authority to direct and take necessary action to insure success."

A most important finding of the three pronged review I mentioned earlier was the realization that changed conditions could out-date our purchase contracts due to a requirement that is not adequately described. I have in mind clause 5e of Standard Contract Form Number 32. This clause reads: "The contractor shall provide and maintain an inspection system acceptable to the government covering the supplies hereunder. Records of all inspection work by the contractor shall be kept complete and available to the government during the performance of this contract and for such longer period as may be specified elsewhere in this contract." The question arises as to how to decide when a contractor's system is acceptable. Should we include all of the procedures and actions taken to produce an acceptable product? If this were done we

would in fact be violating our policy of holding contractors responsible and vesting in them commensurate authority. In order to keep abreast of the times, the Department of the Army has decided that contract clause 5e, to have reasonable meaning, should be implemented with more definiteness. This is accomplished, in cases where definite specifications are available, by agreeing that a contractor's system is acceptable when his supplies consistently meet requirements and when he effectively performs, as a minimum, the quality assurance provisions of the specifications, prior to presentation of supplies for acceptance. In order that contractors may completely evaluate costs, the Army has further established basic policy which holds that the Quality Assurance Section of the specification establishes a limit on the maximum severity of Army quality assurance actions. This basic policy which covers a major portion of Army purchases is further pursued in Army Regulations and reads: "Where there is satisfactory evidence of high quality of production which is the definite result of an effective quality control and inspection system, the amount of government inspection will be adjusted to a minimum consistent with proper assurance that the supplies conform to the quality requirements established by the procurement documents." Also included in the regulations are the general details for implementation of the policy. Since under this program sizable contracts are many times handled by only one or two Army inspectors, some curb on situations that might create surges in the need for Army inspectors is necessary. Therefore, the Army requests contractors to reduce their system to a formal written form and to agree to give prior notice of proposed changes to the system. I want to stress at this point that implementation of the Army's Reduced Inspection Program is voluntary on the part of the contractor.

The procedures I have described apply to all cases where we can describe definitely the item we desire and the Quality Assurance Provisions to be used in determining acceptability. There is another category of items becoming more and more prevalent in military purchasing. These are items like guided missiles, anti-aircraft systems, and similar complex, "on-the-edge-of-our-technology, wonder weapons." Several peculiarities become readily apparent with such items. Seldom are they ever fully defined. Changes, refinements, and significant redesign seems to be the rule, not the exception. To those of you who have worked on our regular equipment this may appear to be normal. The normalcy is lost in this category of items because the changes occur at a fantastic rate not experienced in items which have been stabilized. The reason lies in their newness and their extreme complexity coupled with the urgency of having the item in being. All this leads to inability to prepare firm Quality Assurance Provisions for after-the-fact acceptance inspection. Quality Control from the cradle to the grave is a must on these items. For this the Army has published a specification describing the requirement for a contractor's quality control system. To date, there has been little experience in the use of this or similar specifications on Army items by Army organizations. However, a wealth of experience has been accumulated by our sister department, the Air Force. This experience applies to items that are in many ways similar to the Army items, which when purchased, use the general quality control specification. The Army Inspection Council has a current project to review Air Force experience and to obtain Air Force views and recommendations on the use of quality control specifications. The results of this project will be proposals of adaptations suitable for use by Army organizations on Army items.

Summation

In summation, the rapid growth of new Quality Control concepts has posed problems in keeping up with the times. Army is keeping pace with this rapid growth and is solving problems as they arise with a minimum of disruption of the American way of life. Contractors will still be in charge of their own affairs. To accomplish this goal, the Department of the Army has established a flexible three-fold Quality Control and Inspection program.

First, in a majority of instances, where definitive specifications can be prepared and where industry is well versed in quality control, Army inspection will be much reduced. Benefits are lower costs, to both the Army and industry, smoother relations, and enhanced productivity.

Second, where Quality Control has not been effectively adopted by industry, contractors may obtain guidance for their inspection operations and information of the government's procedures. The results will be, protection of the government's interests at the lowest cost, consistent with the quality offered for acceptance.

And third, where urgency, newness, and complexity militate against definitive item specifications, a Quality Control specification will be used to assure optimum control under the special problems such situations generate.

We believe that the Army's flexible Quality Control and inspection program is of such nature as to provide maximum economy to the government and the contractor and maximum utilization of the contractor's quality control.

CONTRACT REQUIREMENTS AND QUALITY CONTROL

William K. Chormley, Brigadier General, USA
Staff Director, Purchasing & Contracting Policies Division
Office of the Assistant Secretary of Defense
(Supply & Logistics)

Quality is very important to the DOD procurement. Without high quality materiel, certainly in the weapons area, and to only a slightly less extent, in the weapons support area, the Armed Services would not only have less than the most dependable materiel but also would be required to spend needless effort maintaining the materiel, resulting conceivably in failure in battle. Inspection is a technique for assuring that the DOD secures materiel which meets the specified standard, and quality control offers a basis for making the engineering judgment short of an every item, every process inspection.

Procurement is the process utilized in the obtaining of supplies and services which are required to support the military activities of the Nation. These supplies and services are, as a general matter, supplied by industry. Even in those few instances in which the work is performed in a government arsenal, shipyard or maintenance shop, the necessary spare parts and raw materials are procured from industry.

In terms of commodities, the DOD purchases some three million items, including, on the one hand, such simple every day commodities as hand tools, shoes and milk, and on the other hand, tremendously complicated aeronautical and electronic items. In terms of services, the program includes research of the most difficult type, personal and professional services of almost every type, and the like.

Our programs sometimes necessitate unique purchases. A little while ago, my attention was called to a purchase of some leeches which were required for experimental use. The transaction in question resulted in quite a quality discussion - one bidder contending that while the leeches he offered were of the highest standards, being grown in the purest of spring water, leeches offered by the low bidder were grown in "contaminated" waters.

As a matter of interest, the leeches were being purchased for a use which did not require "pure-bred" leeches, the "contaminated" product being quite satisfactory for our intended use. This story serves to demonstrate the fundamental proposition that we purchase to fulfill our needs, and we ought to stipulate that quality level which will fulfill those needs, so long as that quality level is reasonably obtainable.

In terms of scope of procurement, the DOD purchasing involves for 1957, an expenditure of some 20 billion dollars, broken down roughly as follows (in billions of dollars):

Aircraft	6.8	Maintenance and equip	1.5
Guided missiles	2.6	Construction	2.
Ships	1.5	Research and Development	1.6
Electronics	1.	All other procurement	2.0
Petroleum, oil & lubes	1.		

As to the number of transactions, it may be interesting to note that during FY 1956, there were 5.5 million transactions awarded by some 1000 purchasing offices located throughout the world. While the bulk of these purchases involved purchase order or other simplified purchasing methods, the vast preponderance of the money was expended in some 100,000 transactions. Every transaction involved the problem of product description, or the establishment of an appropriate specification or standard, and included a requirement with respect to the specified quality level. Each transaction involved one or more inspections to ascertain compliance with the specified quality level. This inspection was accomplished by approximately 19,000 personnel assigned to the quality control supply and procurement elements of the Army, Navy and the Air Force.

Even though a contract is silent on the subject of inspection, there is a long history of judicial decisions which provide a common-law basis for inspection and quality assurance. Also, in light of the fact that most business conducted within the United States is interstate in character, the industrial associations and concerns have promoted a "Uniform Sales Act", which Act has been adopted by most of the States. This Sales Act includes, of course, many provisions affecting both inspection and acceptance as these terms are defined above. Let us look at a few principles to demonstrate certain aspects of the problem:

(a) Right of Inspection and Right of Rejection.

The courts have held that a buyer is entitled to a fair opportunity to inspect or examine the article or commodity tendered to determine whether it conforms to the contract, that is, whether it is such as was bargained for, and if the article or commodity does not correspond in kind, quality, condition or amount to that he contracted for, the buyer may reject it. Specifically, in relation to the Federal Government as a contracting party, it was held (U.S. v Smoot, 15 Wall (U.S.) 36) that in the complete absence of a contractual provision, the Federal Government had the right to make reasonable rules regulating the inspection of goods it has agreed to purchase, and to prevent the perpetration of fraud on it.

(b) What constitutes Acceptance?

The courts have ruled that there may be several KINDS of acceptances. For example, acceptance of TITLE may be distinguished from acceptance of QUALITY. While acceptance of quality and acceptance of title are usually contemporaneous, there have been cases in which the courts have held that even though title passed, the purchaser may recover damages or rescind the sale if the goods, upon inspection, proved to be of a quality inferior to that required by the contract. In the latter instance, the buyer is also given a reasonable time within which to conduct his quality acceptance. In government procurement, this concept is usually covered specifically by the inclusion of a guarantee, or more properly, a warranty clause in the contract.

The above establishes the fact that in common law, as well as in statutory law, distinctions are drawn between INSPECTION and ACCEPTANCE. In government procurement this distinction is maintained. But, although they are separate considerations, they are closely related. To show the closeness of inspection and acceptance, both concepts are covered in Section XIV of the Armed Services Procurement Regulation;

and to show the separation of the two, separate Parts are provided, Part 1, covering Inspection and Part 2, Acceptance.

The ASPR definitions continue the distinction. Inspection means "the examination (including testing) of supplies and services (including, when appropriate, raw materials components, and intermediate assemblies) to DETERMINE WHETHER THE SUPPLIES OR SERVICES CONFORM TO CONTRACT REQUIREMENTS, which include all applicable drawings, specifications, and purchase descriptions." To coin a phrase, inspection is specification enforcement.

Acceptance, on the other hand, is defined as: "The act of an authorized representative of the Government by which the Government assents to ownership by it of existing and identified supplies... as partial or complete performance of the contract." Acceptance means too that the government contract administrator has indicated compliance with the many other contractual provisions IN ADDITION TO THOSE RELATED TO THE SPECIFICATION.

There are many contract clauses which have an influence upon the quality aspects of the contract:

a. The Inspection Clauses.

The DOD inspection clauses are very simple in construction. In addition to the establishment of an agreed solution in the event of non-conforming supplies, the clauses have two provisions: (i) they require that the contractor "provide and maintain an inspection system" and to maintain the "records of all inspection work...and make them... available to the Government", and (ii) they provide that the material "shall be subject to inspection and test by the Government ...(... including raw materials, components, intermediate assemblies, and end products) to the extent practicable at all times and places including the period of manufacture..."

You will note that the requirements of the clauses are very general in nature, and are suitable for general application to almost any contract or any materiel. No standard or quality levels are stipulated nor test techniques specified. The absence of these in the inspection clause does not mean that they are non-existent, but merely that they are supplied elsewhere. They are supplied as a part of the specification or standard, which, of course, are incorporated by reference into the contract, and thus become contractual as between the parties. In addition, it is to be noted that there is a contractual requirement that the contractor "maintain" an inspective system acceptable to the government, but there is not provided implementing directions with respect to an adequate inspection system.

b. The Changes Clauses.

Another contract provision affecting quality and inspection are the "Changes" clauses. These give to the government the unilateral right to make changes in "drawings, design or specifications, where the supplies to be furnished are to be specially manufactured for the Government in accordance therewith..."

Lest you believe it unfair that the government be given the right unilaterally to amend the specification, it should be borne in mind that the vast preponderance of all changes are "engineering changes" which spring from proposals of the contractors themselves, and that, in any event, the parties have agreed to negotiate and EQUITABLE ADJUSTMENT in the pricing growing out of the change. This situation is justified, in my view, by the basic requirement that the public receive the soundest product possible. I believe, furthermore, that all purchasing organizations are aware of the need for "equity" in the adjustments in prices growing out of the application of the clause. With the assurance of equitable treatment in the background, and in order that the production will continue, provision is made additionally to the effect that the contractor will continue under the "contract as changed." In addition, inability to negotiate a suitable equitable solution to the change, the matter is solved by an appeal to the Armed Services Board of Contract Appeals under the Disputes Clause, next discussed.

c. The Dispute Clause.

All Department of Defense Contracts, as well as all governmental contracts contain a Disputes Clause which offers a quasi-judicial departmental adjudication of unresolved disputes as between the contractor and the contracting officer. The net effect of this clause in DOD contracts is that agreement is reached between the parties that the Armed Services Board of Contract Appeals will hear and determine such unresolved disputes, subject to an appeal to the courts if the decision is "determined by a court of competent jurisdiction to have been fraudulent or capricious or arbitrary or so grossly erroneous as necessarily to imply bad faith, or not supported by substantial evidence." Included within the frame-work of this clause would be problems which grow out of specifications and standards, inspection and quality.

Over a period of years, the Armed Services Board of Contract Appeal has merited the respect of both industry, government and the legal profession for its fairness and equity in the application of legal principles to the disputes at hand.

The independence of the Board, together with the appreciation of industry and the Bar Associations of the fairness of the decisions of the Board, have largely negated an apparent situation in which disputes are determined by one of the disputants - the Government. As a matter of fact, there are very few cases in which contractors seek further relief in the courts.

d. Warranty and Guarantee Clauses.

While there is no guarantee or warranty clause for DOD-wide use, the military departments have utilized such clauses for many years. The clauses are all similar in import: (i) there is a warranty concerning defects in material or workmanship, (ii) there is a warranty that the item conforms to the specification standard or its intended use or both, and (iii) a period is designated during which the warranty is effective, usually six months or a year. Warranty provisions are generally used on a selective basis in those situations in which the government buyer deems that the government needs the additional protection that the clause affords.

While there are additional clauses affecting quality, such as purchase by sample, first item approval, and even the approval of components parts, I would like to conclude with a discussion of some considerations which may lead to a clarification of our present inspection clauses.

You will recall that in the discussion of the inspection clauses it was stated that there is a requirement that the contractor maintain an inspection system acceptable to the government, otherwise undefined, but there is not provided the necessary implementing direction with respect to an adequate inspection system. In short, a general obligation in relation to an inspection system is provided, but no specific contractual standards are imposed upon the contractor in relation to the system.

In the DOD there is an expanded effort to see to it that the contractor undertake the management of his contract, with at least minimum involvement by government procurement personnel in that management. This general emphasis has taken several forms in the past:

a. In the accounting field, there is a growing tendency to recognize that the management of the contractor includes the responsibility to maintain an appropriate accounting system to meet the contractor's needs, and if a governmental cost problem presents itself, to meet also the governmental needs. Accordingly, the present tendency is to provide an accounting standard which represents the governmental needs, inspect the accounting system to see to it that the system, in practice, provides for those needs, and thereafter minimize detailed audits and rely upon audits which reflect continued appropriate application of the approved system.

b. Again, in connection with the furnishing of government property for use in governmental procurement, we have determined as DOD policy to reduce the amount of government record keeping which will be done in connection with the property. This was done with the feeling that the contractor must himself set up a system of accountability for the property, and we will stipulate a system for the care of the property, and administer the system to the extent of assuring application of the approved system.

c. Recently, in connection with the approval of subcontracts, we have again applied the concept of the approval of the Contractor's purchasing system instead of the pre-existing approval of his individual subcontracts.

As I see it, this also is the direction in which our inspection actions will move. This will mean that we will stipulate contractually, standards of inspection and will approve the system for compliance with the standards; and will rely more and more upon the contractor for utilization of his approved system.

As a matter of fact, the Air Force has already gone far in this direction in the publication of Standard MIL-Q-5923C(USAF), entitled QUALITY CONTROL REQUIREMENTS, GENERAL. In this Standard, there is initially stated "the general requirements for the establishment of a quality control system by the contractor to assure that materials, supplies, or services meet the quality standards established by the

contract." The standard, additionally, requires that the contractor maintain a "written", fully "planned and developed" system of quality control "based upon consideration of complexity of design, interchangeability and reliability requirements and manufacturing techniques. The system shall assure that adequate control of quality is maintained throughout the entire process of manufacture..."

Of course, it must be borne in mind that a program of the size and complexity of the DOD procurement program above discussed, will sometimes require continuation of every item, every process inspection. Thus, inspection of the contractor's system will become one of tools in the contract administration kit for such use as may be found appropriate as is also the case in the instances of the accounting system, the government property responsibility obligation, and in the purchasing system area.

During the course of this presentation I have presented a word-picture of the current DOD procurement operation in terms of the size and nature of the purchasing program in (i) classes of items being purchased, (ii) number and types of transactions and (iii) number and distribution of our procurement offices. I have also provided some of the DOD procedures which touch or influence quality and some of the current influences which may result in certain changes in the quality area. In conclusion, I hope that you will find that the procurement quality kit-of-tools is adequate to permit both industry and government to perform their respective roles in the shaping of a stronger America.

DISCOVERY SAMPLING BY ATTRIBUTES - REVIEW, REVISIONS AND RESULTS

Ervin F. Taylor
Westinghouse Electric Corporation
Atomic Fuel Department

Discovery Sampling by attributes, a unique approach to sampling inspection, is -when properly applied- one of the most efficient sampling systems in existence. The technique, developed in 1950, was designed to satisfy the requirements of inspection personnel and to overcome some of the objections to other sampling methods.

Since its initial development, Discovery Sampling has been discussed at quality control conferences and meetings in many parts of the country. Response to the method has been exceptional. The comments expressed most often have been that: (1) Discovery Sampling is not new, but is a statistical expression of "pre-statistical quality control" intuitive sampling plans; and (2) non-formalized versions of Discovery Sampling are being employed in many inspection operations.

Following are: a review of the basic Discovery Sampling technique, illustrations of recently incorporated revisions, and descriptions of some results obtained. In addition, there are recommendations to those interested in applying the method, and to those who wish to expand statistical quality control theory by developing corollary methods.

REVIEW

I - No sampling error or risk is involved when the lot being sampled is either 100% good or bad. Sampling risks occur only when a lot is partially defective.

II - Sampling risks are inefficiently controlled when 100% good and bad lots are considered, since they have no influence on that risk. Only partially defective lots affect the sampling risk.

III - Partially defective lots constitute, generally, less than half of the total lots inspected. The balance consists of 100% good and bad lots.

IV - The fractions defective of partially defective lots possess a relatively stable distribution which indicates that low fractions defective are more likely to occur than large fractions defective (Fig. 1).

V - The number of human inspection errors is directly proportional to the size of the sample being inspected.

These are five points which led to the development of Discovery Sampling and which form the underlying theory. The result is a sampling system with a simplicity which minimizes problems of education, administration, and operation.

The basic steps necessary to effect a Discovery Sampling installation are as follows:

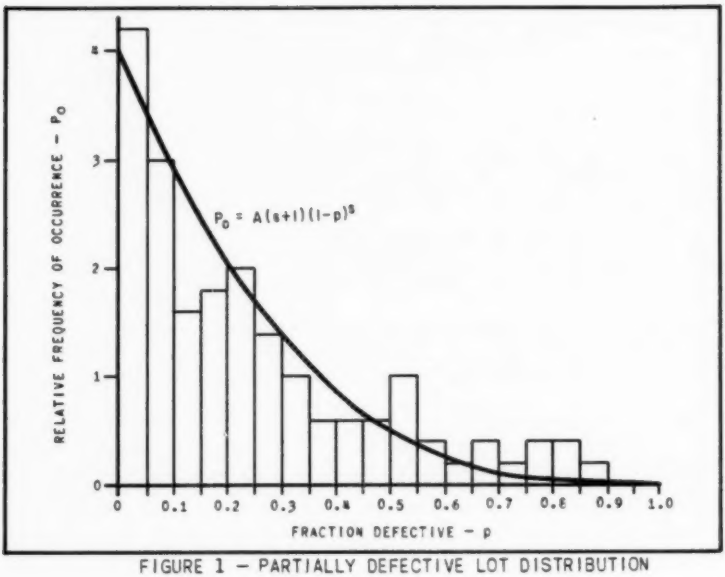


FIGURE 1 – PARTIALLY DEFECTIVE LOT DISTRIBUTION

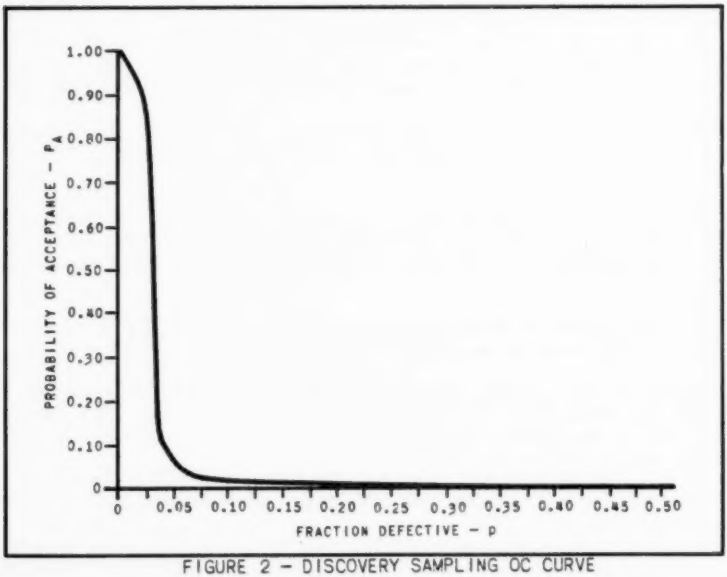


FIGURE 2 – DISCOVERY SAMPLING OC CURVE

1. Make four policy decisions.

   a. Select a target AOQL (average outgoing quality limit).

   b. Select AQL's (average quality levels), if required, for various characteristic classifications (used only in conjunction with the evaluation sample).

   c. Determine the scope of a particular Discovery Sampling plan; i.e. should a specific part number, type of product, entire department's production, etc., be sampled?

   d. Determine the period of sampling control; i.e. should compensations for changing sampling risks be made daily, weekly, monthly, etc.?

2. Collect five items of data from inspection records or special data collection procedures.

   a. The number of partially defective lots inspected.

   b. The number of 100% good lots inspected.

   c. The total number of lots inspected.

   d. The number of lots sample inspected.

   e. The average sample number from lots sample inspected.

3. Plot the frequency distribution of the partially defective lots in 0.05 fraction defective intervals. Compare the distribution with the curve: $P_0 = A(s+1)(1-p)^s$ using several values for s.

4. Compute the Discovery Sampling sample number from the equation:

$$n = \frac{A'}{4(AOQL)} * \frac{(s+n'+1)}{(s+n'+1)-B(s+1)}$$

   where n = new sample number
   n'= old sample number
   A'= reported fraction of partially defective lots
   s = partially defective lots distribution parameter (usually = 3)

Note: If the AOQL = 0.5% and s = 3, the nomographs in Ref. 1 or Ref. 2 will simplify calculations.

5. Establish some regular data collection method for the items in (2) above.

6. Design a flow chart illustrating the steps to be followed. (See Refs. 1 and 2.)

7. Train personnel, using the flow chart. Sampling bowls can be helpful in training.

8. Place the plan in operation; monitoring initially through personal contact with inspectors and their supervisors, and later by control charts ($\bar{X}$ and R) for the sample sizes.

The complete Discovery Sampling theory is given in both references previously cited, and in the Appendix to this paper.

In actual practice instructions to the inspector are quite simple. A typical example would be:

1. Select 8 pieces at random from the lot.

2. Inspect thoroughly all required characteristics of all 8 pieces.

3. If none of the 8 pieces are defective, accept the lot.

4. If any of the 8 pieces are defective, reject the lot.

   4a. Alternative to (4): Retain the sample and select 92 additional pieces (100 total - evaluation sample) at random from the lot.

   4b. Inspect the 92 additional pieces for only those characteristics found defective in the 8 piece sample.

   4c. Accept the lot if less than 3 of any one characteristic are found defective in the entire 100 piece sample.

   4d. Reject the lot as soon as 3 of any one characteristic are found defective in the entire 100 piece sample.

This plan, properly applied, will assure an AOQL of 0.005 through (4) above, an AQL of 1.5% if the alternative to (4) is used, and an ASN (average sample number) of 11.15 pieces per lot.

REVISIONS

Two major revisions have been made to the Discovery Sampling theory since the publication of Ref. 1. These improvements are the OC (operating characteristic) curve calculations and the ASN theory.

Ref. 1 contained OC curve calculations which were rather unconventional and difficult to interpret. A more conventional method was developed and appeared in Ref. 2. An example of the new OC curve for the sampling plan cited above is shown in Fig. 2. (The theory of the OC curve calculations is given in the Appendix to this paper.)

An ASN curve depicting the relationship of the average sample number and the fraction defective of incoming material can be developed for most sampling plans. With this curve and a value for the incoming fraction defective, the average sample number can be determined.

Discovery Sampling is unique in that information is obtained on the frequency of occurrence of partially defective lots. In other words, the proportion of lots likely to be received at any given fraction defective is known. Hence, an average sample number, not a curve, can be

calculated for Discovery Sampling. This takes some of the "ifs" out of the ASN calculation and permits a much more realistic estimation of the inspection work load.

RESULTS

Evidence at ASQC meetings and conferences from coast to coast indicates that most inspection personnel are enthusiastically in favor of Discovery Sampling. In many cases, where a sound application has been made, it has been at the insistence of inspection personnel. The reason is apparent. Most sampling plans require sample sizes which appear unreasonably large to the inspector. In fact, the inspector, who feels that the sample is too large and who is not rigidly controlled, will sample the sample. The good inspector does, however, sample the sample on some rational grounds. He inspects less material from a normally good supplier, and more from a normally questionable supplier. He does the same thing with material from his own shop. He has some knowledge of the quality capabilities of the man-material-machine combination which produces the parts. In actuality, the inspector is establishing an intuitive probability of occurrence of defective lots and adjusts his sample number accordingly. Discovery Sampling does the same thing that most good inspectors are doing, but on a more scientific basis, yielding sample numbers generally much smaller than those usually encountered. This leads to another axiom: Better quality control is effected through a thorough inspection of a small sample than a cursory inspection of a large sample. These are the primary reasons why the technique is accepted so readily by inspectors and inspection supervisors.

One of the best examples of a realistic test to learn the value of Discovery Sampling can be shown by citing the results of Company X. This company had been using MIL-STD-105A for some time. It was decided to test Discovery Sampling with a minimum of disturbance to the then current inspection method.

An AOQL of 0.5% was selected for Discovery Sampling and an AQL of 0.65%, inspection level II-normal, was already being employed for MIL-STD-105A. The Discovery Sampling sample number was calculated at 6 for bolts and 9 for nuts, the products selected for comparison. The former was used regardless of the lot size; the latter was doubled when the lot size exceeded 1000.

The technique used was to select the MIL-STD-105A sample at random from the lot, the first 6 (9 or 18) pieces being the Discovery Sampling sample. A tabulation of the results is given in Table I.

A total of 325 lots of bolts (9 to 121,185 pieces per lot) were compared. Discovery Sampling duplicated the MIL-STD-105A results 98.2% of the time with only 7.1% of the inspection effort. At the 5% level, there was no significant difference between the process average estimates of the two plans.

A total of 202 lots of nuts (19 to 52,300 pieces per lot) were compared. Again, Discovery Sampling duplicated the MIL-STD-105A results 96.0% of the time with only 14.7% of the inspection effort. The increased amount of Discovery Sampling inspection compared with that for bolts can be partially explained by the fact that more than half the lots were over 1000 pieces and required the 18 piece sample.

TABLE I – TEST RESULTS OF DISCOVERY SAMPLING (A) COMPARED WITH MIL-STD-105A (B)

	SAMPLING PLAN	NUMBER OF LOTS	NUMBER OF PIECES IN LOTS	NUMBER OF PIECES INSPECTED	NUMBER OF PIECES DEFECTIVE	ESTIMATED PROCESS AVERAGE	NUMBER OF LOTS ACCEPTED BY BOTH A AND B	NUMBER OF LOTS REJECTED BY BOTH A AND B	NUMBER OF LOTS ACCEPTED BY B REJECTED BY A	NUMBER OF LOTS REJECTED BY B ACCEPTED BY A
TEST 1 – BOLTS	A	325	(2) 856,402	1,934	94	(3) 4.9%	299	(4) 20	2	4
	B			27,214	1,496	5.5%				
	DISCOVERY SAMPLING DUPLICATED MIL-STD-105A (8) 98.2% OF THE TIME WITH ONLY 7.1% OF THE INSPECTION EFFORT.									
TEST 2 – NUTS	A	202	(6) 713,814	2,745	306	(3) 11.1%	150	(7) 44	5	3
	B			18,656	2,040	10.9%				
	DISCOVERY SAMPLING DUPLICATED MIL-STD-105A (8) 96.0% OF THE TIME WITH ONLY 14.7% OF THE INSPECTION EFFORT.									
TESTS 1 & 2	A	527	1,570,216	4,683	400	(3) 8.5%	449	64	7	7
	B			45,870	3,536	7.7%				
	DISCOVERY SAMPLING DUPLICATED MIL-STD-105A (8) 97.3% OF THE TIME WITH ONLY 10.2% OF THE INSPECTION EFFORT.									

NOTES:
(1) CONSTANT SAMPLE NUMBER.
(2) RANGE OF LOT SIZES 9 TO 121,185.
(3) DIFFERENCE NOT SIGNIFICANT AT 5% LEVEL.
(4) INCLUDES 14 100% DEFECTIVE SAMPLES WITH BOTH A AND B.
(5) CONSTANT SAMPLE NUMBER DOUBLED FOR LOTS OVER 1000 PIECES.
(6) RANGE OF LOT SIZES 19 TO 52,300.
(7) INCLUDES 5 100% DEFECTIVE SAMPLES WITH A ONLY AND 19 WITH BOTH A AND B.
(8) ASSUMES MIL-STD-105A AS THE STANDARD.

RECOMMENDATIONS

Discovery Sampling is not being touted as a panacea for all the inspection sampling problems. It does, however, have many excellent applications. When Discovery Sampling cannot be applied, it will make this indication. By varying the sample size according to the percentage of partially defective lots inspected, and maintaining an $\overline{X}$ and R chart of the resultant sample numbers, it soon becomes evident whether or not Discovery Sampling is properly applied. Such a control chart remaining in control after some optimum operation period will indicate that the technique is satisfactory. A chart out of control indicates that some investigation should be made. A study of the situation can often reveal improper or inadequate specifications, a particularly bad supplier or part, or even improper inspection techniques. Once the situation is corrected, Discovery Sampling can probably be applied. When Discovery Sampling fails (only once in the author's experience) it is probably due to the process and not the sampling plan.

The control chart test (if performed as indicated in the preceding paragraph) can be made quite satisfactorily with a minimum of effort and without disrupting the normal inspection sampling operations.

Discovery Sampling can be tested, as indicated in the preceding section, with little disturbance of the existing inspection method. The following procedures are suggested:

1. Follow the 8 basic steps previously outlined.

2. Select the Discovery Sample from the first n pieces of the usual inspection sampling method. (Select at random from lot in case of 100% inspection.)

3. Inspect sample and record results.

4. If a defective is discovered, select additional pieces from those selected for the usual sampling method until either a number of defectives equal to the rejection number at the selected AQL (Table III) is discovered or until a total of 100 pieces has been selected. (Additional pieces for the evaluation sample may have to be randomly selected from the lot if the usual sampling plan calls for less than 100 pieces.)

5. Inspect evaluation sample and record results.

6. Inspect balance of pieces in usual sampling plan and record results.

7. It will be more convenient if a data sheet is established with the following column headings:

   a. Lot number.

   b. Lot size.

   c. Usual inspection plan sample number.

   d. Defectives found in (c) above. Note: More than two columns for (c) and (d) may be required if double or

multiple sampling is currently used.

e. Discovery Sampling sample number.

f. Defectives found in (e) above.

g. Additional evaluation sample pieces inspected.

h. Defectives found in (g) above.

i. Lot accepted by both plans.

j. Lot rejected by both plans.

k. Lot rejected by Discovery Sampling; accepted by other plan.

l. Lot accepted by Discovery Sampling; rejected by other plan.

This test, in addition to the control charts of the Discovery Sampling sample number, should provide ample evidence on the value of the technique.

Practically the only criticism made of Discovery Sampling by informed statisticians has been directed at the empirical partially defective lot distribution curve, $P_0 = A(s+1)(1-p)^s$. The criticism has been that the curve is not theoretically valid and is probably not stable. Since this is an empirical equation, the only answer to such criticism is to suggest the collection of data to prove or disprove its rationality.

The following recommendation is made to any statistician who wishes to perform an economic service to inspection departments everywhere:

Discovery Sampling has introduced to quality control theory the use of a third probability in sampling -- the probability of occurrence of defective lots in addition to the probability of a defective piece in a lot (p, the fraction defective) and the probability of acceptance (P_a). The distribution as described by Discovery Sampling may not be theoretically perfect. It is reasonable, however, that such a distribution or distributions do exist. It is sincerely recommended that interested statisticians develop more theoretical bases for the probability of occurrence distributions. Work of this type would make an extremely valuable contribution to the science of quality control.

CONCLUSION

A number of refinements in Discovery Sampling have been made since its introduction. Many quality control people are talking about the technique, testing it, and applying it to inspection operations. Through Discovery Sampling, many inspection dollars are being saved each year with a closer known control of quality.

The strongest recommendation which can be made for Discovery Sampling is: It works.

APPENDIX

Note: The complete theory of Discovery Sampling is included in this Appendix to place the OC curve and ASN theories in their proper perspective.

BASIC THEORY

Discovery Sampling considers the probability of occurrence of a partially defective lot, the actual distribution of partially defective lots; as well as the probability that a sampling plan will accept the lot. Consideration of these probabilities determines an average outgoing quality limit and facilitates the construction of OC curves where the probability of occurrence of a partially defective lot is considered.

Lots presented for acceptance fall into three mutually exclusive classes.

Class	Symbol	Fraction Defective
100% Effective	(GDL)	$p = 0$
Partially Defective	(PDL)	$0 < p < 1$
100% Defective	(none)	$p = 1$

A lot which is 100% defective does not constitute a sampling risk; since it will be discovered if only one item is inspected. Hence, these lots will be excluded from further consideration.

The probability that a partially defective lot will occur is defined as: the ratio of the number of partially defective lots to the number of partially defective lots plus the number of 100% effective lots which are presented for acceptance during a given interval of time. Symbolically:

$$A = \frac{\Sigma_t(PDL)}{\Sigma_t(PDL)+\Sigma_t(GDL)} \qquad (1)$$

A study was made to determine the distribution of partially defective lots. The probability density function

$$f(p;s) = (s+1)(1-p)^s dp; \quad s \geq 0, \ 0<p<1 \qquad (2)$$

was found to represent this data on a conservative basis. The value of the parameter "s" determined from the data was approximately 3.*

The probability of occurrence of a partially defective lot with fraction defective p may now be defined as:

$$P_a = A(s+1)(1-p)^s dp \qquad (3)$$

The probability that a lot with fraction defective p will be accepted by a sample of size "n" with no defectives allowed is approximately:

$$P_a = (1-p)^n$$

* The question arises, would other studies also give this distribution? The writer has made many studies of partially defective lots. In every case Eq. (2) was applicable, although at times very conservatively.

Therefore, the probability that a lot with fraction defective p will occur and be accepted is:

$$P_0 P_a = A(s+1)(1-p)^{s+n}dp \qquad (4)$$

Assume that for each lot of size k there exists a set of lots of size k which have the distribution defined by Eq. (2). Then the following is true for each set of lots and hence for all sets of lots.

A lot with fraction defective p contributes to outgoing quality the fraction defective:

$$FD = A(s+1)p(1-p)^{s+n}dp$$

And the total fraction defective contributed to outgoing quality for all partially defective lots is:

$$\Sigma pFD = A(s+1)\int_0^1 p(1-p)^{s+n}dp = A\frac{s+1}{(s+n+1)(s+n+2)} \qquad (5)$$

(Screening of lots in which defectives are found is assumed.)

Similarly the total fraction effective contributed to outgoing quality for all partially defective lots is:

$$\Sigma pFE = A\left[1-(s+1)\int_0^1 (1-p)^s pdp\right] = A\frac{s+1}{s+2}$$

Also the total fraction effective contributed to outgoing quality for all 100% effective lots is:

$$\Sigma FE = 1-A$$

The average outgoing quality may now be defined as:

$$AOQ = \frac{\Sigma pFD}{\Sigma FE+\Sigma pFE+\Sigma pFD}$$

or

$$AOQ = \frac{A(s+1)(s+2)}{(s+2-A)(s+n+1)(s+n+2)+A(s+1)(s+2)}$$

Considering AOQ as a function of s, it is found that AOQ has a maximum value. Thus, we may define average outgoing quality limit (AOQL) as:

$$AOQL = \frac{A}{4n}, \quad A \le \frac{1}{2} \qquad (6)$$

In Eq. (1) it was assumed that the true value of A was known. However, if sampling was applied, this is not the case. Let "B" denote the fraction of lots to which sampling was applied. Then from Eq. (4) it follows that:

$$BA(s+1)\int_0^1 (1-p)^{s+n'}dp = BA\frac{s+1}{s+n'+1}$$

(where n' = size of sample actually used.)

is the fraction of partially defective lots which would have been considered 100% effective lots. Thus if A' is the reported value of A then:

$$A' = A - AB\frac{s+1}{s+n'+1}$$

Therefore Eq. (6) can be written as:

$$n = \frac{A'}{4(AOQL)} \cdot \frac{(s+n'+1)}{(s+n'+1)-B(s+1)} \qquad (7)$$

From Eq. (7) when n', B, A' and s are known, n can be determined in order to ensure any AOQL.

OC CURVES

From Eq. (6), it is evident that if the AOQL is fixed, then for each value of A there is a corresponding n. This relation is shown in Table II with AOQL = 0.5%.

In order to avoid screening of lots a second sample of 100-n items is taken if the first sample contains any defectives. From Eq. 4, it follows that the probability of accepting a lot with fraction defective p on the first sample is:

$$P_A = A(s+1)\int_{p'}^{p'+\Delta p'} (1-p)^{s+n}\, dp \qquad p' \le p \le p'+\Delta p'$$

If AQL's of 1.5% and 4.0% are used (Table III), then the probability of acceptance on the second sample is:

P_M = Probability of obtaining two or less defectives in a sample of 100-n (AQL = 1.5%)

P_m = Probability of obtaining six or less defectives in a sample of 100-n (AQL = 4.0%)

$$P_M = \sum_{k=0}^{1} C_k^{100-n} (1-p)^{100-n-k} p^k$$

$$P_m = \sum_{k=0}^{5} C_k^{100-n} (1-p)^{100-n-k} p^k$$

with $p = p' + \frac{1}{2}\Delta p'$

Therefore the probability of acceptance on the combined samples is:

$$P_{AM} = P_A + P_M \quad (AQL = 1.5\%)$$

$$P_{Am} = P_A + P_m \quad (AQL = 4.0\%)$$

TABLE III

RELATIONSHIP BETWEEN DEFECTIVES ALLOWED PER 100 AND AQL

DEFECTIVES ALLOWED PER 100	AQL - PRODUCER'S RISK 0.05
0	0.1%
1	0.3
2	0.7
3	1.3
4	2.0
5	2.6
6	3.3
7	4.0
8	4.8
9	5.5
10	6.2
11	7.0
12	7.8
13	8.6
14	9.5
15	10.3

COMPUTED FROM REFERENCE (3)

These values are shown in Table II for s = 3, p' = 0.0, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50 with $\Delta p' = 0.05$.

AVERAGE SAMPLE NUMBERS

From Eq.'s (3) and (4) it follows that the probability of no decision P_{ND} on the first sample is:

$$P_{ND} = A(s+1)\int_{p'}^{p'+\Delta p'} (1-p)^{s}\, dp - A(s+1)\int_{p'}^{p'+\Delta p'} (1-p)^{s+n}\, dp$$

TABLE II

PROBABILITY OF ACCEPTANCE AQL = 4.0%											
N	A	p' + Δ p'									ASN
		.025	.075	.125	.175	.225	.275	.325	.375	.425	
1	.02	.984	.003	.002	.002	.001	.001	.001	.001	.000	1.03
2	.04	.967	.005	.004	.003	.002	.002	.001	.001	.000	2.44
3	.06	.951	.010	.005	.004	.003	.002	.001	.001	.000	3.77
4	.08	.934	.011	.006	.004	.003	.002	.001	.001	.000	5.27
5	.10	.918	.013	.007	.004	.003	.002	.001	.000	.000	6.73
6	.12	.903	.015	.007	.004	.002	.001	.001	.000	.000	8.16
7	.14	.887	.017	.007	.004	.002	.001	.001	.000	.000	9.72
8	.16	.872	.020	.007	.004	.002	.001	.000	.000	.000	11.15
9	.18	.855	.021	.007	.004	.002	.001	.000	.000	.000	12.05
10	.20	.840	.024	.007	.003	.001	.001	.000	.000	.000	13.92
11	.22	.823	.025	.007	.003	.001	.001	.000	.000	.000	15.43
12	.24	.809	.027	.007	.003	.001	.000	.000	.000	.000	16.84
13	.26	.794	.029	.007	.002	.001	.000	.000	.000	.000	18.29
14	.28	.777	.031	.007	.002	.001	.000	.000	.000	.000	19.59
15	.30	.761	.033	.007	.002	.001	.000	.000	.000	.000	20.79
16	.32	.746	.036	.006	.002	.001	.000	.000	.000	.000	22.16
17	.34	.730	.038	.006	.002	.000	.000	.000	.000	.000	23.28
18	.36	.715	.041	.006	.001	.000	.000	.000	.000	.000	24.27
19	.38	.701	.042	.006	.001	.000	.000	.000	.000	.000	25.37
20	.40	.685	.045	.006	.001	.000	.000	.000	.000	.000	26.39

PROBABILITY OF ACCEPTANCE AQL = 1.5%											
N	A	p' + Δ p'									ASN
		.025	.075	.125	.175	.225	.275	.325	.375	.425	
1	.02	.984	.003	.002	.002	.001	.001	.001	.001	.000	1.03
2	.04	.967	.005	.004	.003	.002	.002	.001	.001	.000	2.18
3	.06	.950	.008	.005	.004	.003	.002	.001	.001	.000	3.34
4	.08	.933	.009	.006	.004	.003	.002	.001	.001	.000	4.55
5	.10	.916	.011	.007	.004	.003	.002	.001	.000	.000	5.68
6	.12	.900	.012	.007	.004	.002	.002	.001	.000	.000	6.84
7	.14	.883	.013	.007	.004	.002	.002	.001	.000	.000	7.95
8	.16	.867	.014	.007	.004	.002	.002	.000	.000	.000	9.14
9	.18	.849	.014	.007	.004	.002	.002	.000	.000	.000	10.04
10	.20	.832	.015	.007	.003	.001	.001	.000	.000	.000	11.22
11	.22	.814	.015	.007	.003	.001	.001	.000	.000	.000	12.36
12	.24	.798	.015	.007	.003	.001	.000	.000	.000	.000	13.42
13	.26	.781	.015	.006	.002	.001	.000	.000	.000	.000	14.51
14	.28	.763	.015	.006	.002	.001	.000	.000	.000	.000	15.36
15	.30	.745	.015	.006	.002	.001	.000	.000	.000	.000	16.32
16	.32	.728	.015	.005	.002	.001	.000	.000	.000	.000	17.34
17	.34	.716	.016	.005	.002	.000	.000	.000	.000	.000	17.53
18	.36	.693	.016	.005	.001	.000	.000	.000	.000	.000	18.95
19	.38	.676	.015	.004	.001	.000	.000	.000	.000	.000	19.81
20	.40	.658	.015	.004	.001	.000	.000	.000	.000	.000	20.60

Hence the average sample number (ASN) may be defined:

$$ASN_M = n\,P_A + P_{ND}\,ASN'_M \quad (AQL = 1.5\%) \quad (8)$$

$$ASN_m = n\,P_A + P_{ND}\,ASN'_m \quad (AQL = 4.0\%) \quad (9)$$

Where ASN'_M (ASN'_m) is the ASN for a sample of 100-n with 2 (6) defectives allowed in this sample.

It can be shown that[4]:

$$ASN_M = \frac{3}{p} - \frac{3(1-p)}{p}\,P_{3:100-n} + (n-6)\,P_{2:100-n} + \frac{(n-2)p}{(1-p)}\,P_{1:100-n}$$

$$ASN_m = \frac{7}{p} - \frac{7(1-p)}{p}\,P_{7:100-n} + (n-13)\,P_{6:100-n} + \frac{(n-6)p}{(1-p)}\,P_{5:100-n}$$

with $p = p' + \frac{1}{2}\Delta p'$ and $P_{k:r}$ = probability of obtaining exactly k defectives in a sample of r.

Since all of the quantities in Eq.'s (8) and (9) are known these equations can be solved determining an ASN for each p' and $\Delta p'$. If thsse ASN's are summed over $0 \leq p \leq 1$ according to the above divisions (Note: The division 0.5 to 1.0 with a corresponding $\Delta p' = 0.5$ is included), then there results an ASN for the entire plan. These ASN's are shown in Table II.

LARGE LOT APPLICATIONS

The basic theory of Discovery Sampling is independent of the lot size. However, it is advantageous to decrease the probability of accepting an unusually large lot with a large fraction defective; and thus decrease the possible fluctuation in the AOQL.

In order to accomplish this and to maintain the simplicity of the sampling plan, it was decided that a lot with more than 1000 pieces or more than twice the usual number of pieces would be regarded as a large lot. The probability of acceptance of such a lot would be decreased by simply doubling the normal sample size.

SAMPLE NUMBER CONTROL CHART

If no assignable causes for fluctions of A (Eq. 1) are present, then values of A may be assumed to be normally distributed. Since n bears a linear relationship to A (Eq. 6), values of n may also be assumed to be normally distributed.

Control charts for individuals or averages, and ranges can be developed with n as the variable. The usual indications of out-of-control conditions or tendencies will signal the presence of assignable causes and the need for an investigation.

REFERENCES

1. Taylor, Ervin F., "Discovery Sampling"; Proceedings of the Ninth Annual A.S.Q.C. Convention; New York, N. Y., 1955.

2. Taylor, Ervin F., Discovery Sampling by Attributes, North American Aviation, Inc., Columbus, Ohio, 1956.

3. Romig, Harry C., 50-100 Binomial Tables, John Wiley & Sons, Inc., New York, N. Y., 1953.

4. Statistical Research Group, Columbia University, Sampling Inspection, McGraw-Hill Book Company, Inc., New York, N. Y., 1948, pp. 189-214.

AUTOMATIC GAGING OF COMPLEX PARTS BY TAPE CONTROLS, ETC.

W. S. Tandler
The Warner & Swasey Research Corporation

There are times in the history of scientific and industrial development at which the opinion prevails that a peak of development has been reached and that things will now stay the way they are and for a long time. While this opinion prevails among a large number of people, there are others at work who for some reason or another prepare new things which after a while tend to replace the old ones. This causes a struggle during which certain of the new things win out and certain of the old ones remain. During such a period of struggle one should be cautious to predict the outcome, but one should keep abreast with the facts. We live in such a period in which the enormous technological achievements of the late 19th and the early 20th century are now attacked by the new scientific discoveries and inventions whose origins can be traced back to just about the time when the technology of the 19th century and the early part of this century was in full blossom.

The effects of this struggle express themselves in many different ways in different fields but common underlying trends are clearly recognizable. Among those trends, and I believe perhaps the strongest, is the tendency towards application of precise measurements and mathematical expression.

Our field, "automatic gaging of complex parts," is but a very minute portion of the total picture. Still, in it are active the same trends toward increased accuracy and precise mathematical expression.

Our field, specialized as it may be, is still larger than could be dealt with with any degree of satisfaction in a short paper and it appears preferable to concentrate on the smallest complex within this field which still has general significance, general interest, and, I believe, quite general application.

We are concerning ourselves here with the gaging, i.e. the dimensional measurement of small and medium-sized mechanical parts; things like screws, turbine blades, gears, castings, etc. We are not talking about microscopic particles, nor are we talking of buildings, bridges, etc. Within this group are many parts which require practically no dimensional inspection, and others which are of such simplicity that one or two measurements completely define the part. There naturally have been developed methods for the gaging of such parts. Our concern is with those parts that have a large number of surfaces and dimensions; some of which may be independent of each other, others may be closely correlated by tight tolerances. The greater the number of dimensions to be gaged, the closer the tolerances, and the more complex the relationship between dimensions (curved, angular, etc.), the more difficult of course becomes the gaging problem. There is hardly any part that could not be measured today with conventional techniques and whose dimensions couldn't be recorded, interpreted and understood. However, those working in the field have found out that, while the above statement is undoubtedly true, the practical difficulties, and perhaps still more important, the time and cost involved, may make the measuring operation which is theoretically possible, for practical reasons a near impossibility.

The question would remain a theoretical one if it weren't that modern

machinery contains numerous parts which are of extreme complexity and designed to the closest tolerances that our highly developed machining industry is capable of maintaining. A hundred or a thousand measurements on one part are not unusual, and it is known that even ten measurements of great accuracy and complex relations may cause severe problems of instrumentation, measuring operation and interpretation of the measuring results. These requirements and the necessity of having to gage complex parts economically and at a fairly rapid rate, has led to the development of a number of high-precision measuring techniques which are at least capable of handling the accuracy problem and which operate at a reasonably high rate. However, it was found out that the obtaining of measurements, even though at a rapid rate, still leaves something important to be desired, and that is the interpretation and the evaluation of the measuring results.

In a complex (multi-dimensional) part the individual measurement may be of little significance. It is the correlation of all measurements which determines acceptability of a part. It is therefore necessary to correlate numerous individual measurements and to come up with a total answer. This answer is often not just a matter of addition and subtraction but may entail higher orders of computation. The difficulty of such computation, combined with the error possibilities and probabilities both in the part and in the measuring instrument and technique, have often led to a situation in which it is impossible to arrive at a definitive conclusion. The complexity and the resulting confusion is sometimes such that there is either no meeting of the mind or repeated measurements and interpretations vary so much from each other that it is impossible to say what a true statement would have to be. This is by no means an isolated case; it is a serious problem.

I believe it is in this region in which the conventional or traditional technique has found its limitation and where the struggle for application of new principles seems to favor most decidedly the new techniques.

In essence, the new technique must be capable of correlating measuring results, computing them, and coming up with a unique answer.

There are two approaches possible: One is what may be called the graphic technique; the other one, automatic computation which in turn may result in numerical presentation or in automatic assorting or segregation.

By graphic technique is meant an approach by which the measuring results are presented not as a hodgepodge of individual measuring data - be it in form of a table or otherwise - but in the form of a graph or chart which is laid out and designed in such a way that a human operator is capable of interpreting and drawing conclusions from the study of such a graph, and can do this within a reasonable time. Gaging machines of this type have been on the market for approximately twenty years and the trend towards graphic presentation is increasing. It permits the operator to see at one glance what would take sometimes an hour of painstaking computation to evaluate.

Automatic computing of measuring results, in its simplest form, is the "go - no go" gage. Gages of this type have been known for a long time, as applied to individual dimensions. But the technique of combining several measurements in one single computing operation is new.

We employ in our work and in our products both techniques, the chart or graphic presentation and the automatic computation.

Needless to say that a chart can be drawn up from individual measurements by hand and a calculating or computing machine can be used to evaluate a group of individual measurements. Such a procedure is of course old, but the incorporation of the recording system into an automatically operating measuring machine and the incorporation of a computing system into an automatic inspection machine, are of recent origin.

We produce a machine, the Automatic AIRFOIL PROBOGRAPH, which combines both techniques. It first measures automatically, computes the results, and decides whether to accept or reject in accordance with set tolerances incorporated into the machine. If the decision is to accept, no further action is taken. However, if the decision is to reject, the Automatic AIRFOIL PROBOGRAPH, automatically and on its own, records all measurements in the form of a deviation chart so that this chart can serve as a rejection report without requiring further elaboration.

It would not have been possible to build such automatic inspection machinery without applying new measuring techniques as such. I developed over a number of years a technique, known as the PROBOGRAPH method, in which measurements are taken by electric contact (not by spark) and correlated by precision slides, linear-rectangular or rotary-polar. This technique is not only free of the age-old problem of pressure distortion, but it lends itself to automatic operation -- in fact, it is an automatic technique. It did not originate by improving a hand-operated method; it was conceived and designed for automatic operation. It is therefore completely under automatic control. In some instances this control is sequence control, in others program and tape control.

In order to establish, introduce and successfully start the PROBOGRAPH technique, a basic concept of part-design had to go along with it. This concept is coordinate dimensioning. In the early days of the PROBOGRAPH the possibility of coordinate dimensioning of drawings was considered worse than science fiction. It lacked entertainment value. Today coordinate dimensioning is a strong factor, though not yet universally applied. Coordinate dimensioning is not only a basic requirement of automatic measurement of complex parts; it is a basic requirement of automatic computation. The PROBOGRAPH measuring technique and the automatic computor go hand in hand. The computor furnishes data for the PROBOGRAPH to measure, and the data of the PROBOGRAPH can be processed through a computor. Naturally, the use of an intricate large-size computor is only required in exceptional cases. Most of present-day inspection technique revolves around parts which are simple enough to be handled without extensive computing, but which are by far too complex to be handled without a simple and rapid computing mechanism incorporated into the standard PROBOGRAPH models. The extent of computation requirements depends on the application, but the underlying idea of digital computor and of PROBOGRAPH coordinate automatic measuring technique is the same.

This idea, as initially remarked, consists basically of the two elements: higher accuracy, and increased use of mathematical thinking. My own field is of course the application of the PROBOGRAPH technique to the numerous measuring problems which are brought to us or which we attempt to discover and point out. Some of these problems are of a nature in which thousands of similar parts are processed, and in this case not only is speed of the essence,but it is also possible to invest

in a fairly elaborate setup which is called "Tooling" and which permits to take as many as desired, but usually 40 - 50 measurements simultaneously and to come up with an answer within a matter of seconds, and a recorded chart if the answer is rejection. These measuring machines are called multi-probe PROBOGRAPHS.

There is, however, an increasing number of measuring problems in which only relatively small quantities or even single parts have to be inspected and where our technique as such, because of its element of pressurelessness and automatic operation, is of great advantage but where a setup or tooling is for economical reasons or time limitations not practical. Therefore, instead of relying on tooling and taking all measurements simultaneously, we have developed the single-probe PROBOGRAPH in which one (occasionally two or more) probe performs all the measuring operations consecutively instead of simultaneously.

The measuring operations have to be programmed. We have built such programmed measuring machines for many years, and in more recent years these machines were delivered and are being operated with punched tape controls. The punched tape (preferably 8-hole Flexowriter) furnishes the machine not only with the program of its operation, but also with the digital numerical data required. These data fall into two groups: (1) The positioning data, the points at which the measurements are taken; and (2) the measuring data, i.e. the dimensions which the instrument should read at a given position. The PROBOGRAPH, besides automatically positioning, also operates with the nominal dimensions of the part and in its operation automatically compares the nominal with the actual dimensions. The difference between nominal and actual dimension is of course the "error." Dependent on the machine, this "error" is recorded or used for computation, or both.

The techniques of automatically controlling the positioning as well as the indication of nominal dimensions were developed over many years, and as the requirements for more and more complex parts increased, and as the number of necessary measurements became larger and larger, we introduced tape control to an increasing degree, because tape is one of the best means of storing large amounts of data in digital form for the operation of automatic machines. It seems needless to point out that in order to do the required job our system including tape controls had to be not only automatic but also practically unlimited in accuracy as well as definition, that it further had to be absolutely reliable and neither limited in distance of travel nor in operating speed. Our system incorporated in the PROBOGRAPH, PROBOMAT, TELE-PROBOMAT and PILOT PROBOMAT fulfills all these specifications. With such a method it is therefore possible to handle measuring problems of the most complex kind, not only by elaborate tooling, but on a single-piece basis by tape controls.

This development of a fully-automatic tape control positioning and measuring system led to a most significant consequence.

The same controls that are applicable to positioning for measuring are also applicable to machine tool controls in general. We are actually operating with the same system not only our measuring machines, but also jig borers. This fact is of major significance because it proves that the same basic principles of coordinate dimensioning apply both to the making and to the inspection, and if properly applied, can be handled by tape control.

A tape prepared for the automatic control of a jig borer, for instance, can be converted into another tape for use of an automatic measuring machine. The two tapes contain the same basic information, both command and digital type. There are only minor differences due to the nature of the machine process and the measuring process. The conversion from one tape to the other requires only a fixed set of rules and not the thinking and decision of a human operator. Consequently, a part can be produced, for instance on a jig borer, under full tape control. The tape used can be converted for inspection purposes and this new tape serves to inspect the part. It is of course understood that the tape must be of such a nature that it can be properly verified, both for manufacture and for inspection. What is true for the jig borer is also true for practically any other machine tool. The inspection machine need not incorporate the majority of functions of the machine tool; it only must have the same coordinate system built into it in the form of slides and rotary devices. That means, that one universal inspection machine with a proper coordinate system in a sufficient number of degrees of freedom,can be used to check and inspect an almost infinite variety of products. In practical application the following considerations and possibilities are of interest:

1. The measuring machine, though having the same coordinate system as the machine tool, has none of the following problems: no high speed, no tool pressure, no coolants, no chips, no heat generated in the machining process. Therefore, although the control tape may be essentially the same, the measuring machine shows the purified results and indicates the deviations caused by the elements enumerated above. In this manner the nature of these elements can be ascertained and incorporated into the machine tool tape in the form of correction factors, if and where needed.

2. In certain industries, mostly of a military kind, subcontracting is not only economically advantageous but strategically necessary for dispersion. In this case it is possible to establish a relationship between subcontractor using tape for manufacture, and manufacturer using the tape to ascertain that the part delivered corresponds to the specifications without the need of discussions, interpretations and all the ensuing misunderstandings and conflicts.

3. Experience has shown that the introduction of a rational method based on coordinates and automatic controls demands a precision of thinking and specifications which, while sometimes not easy to obtain, have always proven to be of a very healthy influence and have solved many a knotty problem.

One purpose of this paper appeared to me to acquaint you with the origin and the basic problems and implications of automatic gaging of complex parts and how its principles tie in with other fields. We have only made a start, and the biggest results are yet to come. A few illustrations here-attached may help in visualizing the practical applications.

The number and variety of "complex parts" is such that an enumeration of a few of them is all that we can attempt to give here. They form essentially two categories: (1) contoured parts, and (2) non-contoured parts. Among the contoured parts which are of importance are cams, gears, airfoil and waterfoil shapes such as turbine blades, propellers, etc. Most of these parts have also non-contoured surfaces, flats, cylindrical forms, and holes.

Among the non-contoured "complex" parts are housings, such as gear housings, turbine, compressor and pump housings, shafts, and an endless variety of other parts. The close control and inspection of concentricity and alignment of bores, spacing of holes and their relation to other surfaces on the same piece appears to be one of the most exacting jobs at the present moment.

The control of manufacture and inspection of the second category (the non-contoured parts) is at least as great a field as that of the first category (contoured parts). However, contoured parts having suddenly grown into prominence with the advent of gas turbines and high-precision cams, as well as complex gearing, industry was less prepared to cope with these problems and therefore the new techniques such as the PROBOGRAPH found an open door. Therefore, the initial progress of automatic inspection methods of complex parts including tape-controlled measuring machines was made in the field of contoured parts. Our illustrations show a number of these machines. Actually it has become evident today that these same methods that were developed for complex contoured parts can be applied most profitably to non-contoured parts, both in automatic tape-controlled manufacture and in automatic inspection.

Fig. 1

Illustration 1 shows a PROBOMAT tape-controlled automatic inspection machine in operation. To the left of the operator is the measuring unit. The part under inspection is an impeller wheel set up for the measurement of the curvature of its vanes. To the right is the control unit which includes both the tape controls and the recording unit. Measuring and control units are connected by electric cable.

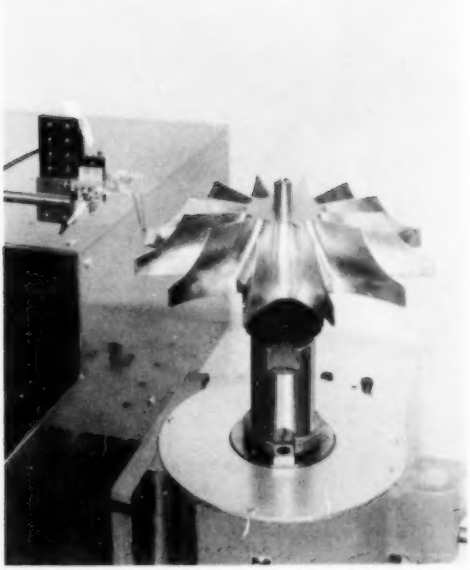


Fig. 2

Illustration 2 shows a close-up of the impeller wheel being measured on the PROBOMAT and the relative position of the probe holder to the vanes of an impeller wheel. The probe holder is mounted on an arm which extends from the measuring head. The probe holder is moved by program control to the positions and measuring points indicated by the print. The wheel is held in a rotary fixture.

Illustration 3 is a detailed view of the probes showing the adjustable probe holder; the arrangement of the probe points in form of a "T". The right-hand probe is measuring a portion of an impeller wheel.

Illustration 4 is a close-up of the control unit of the PROBOMAT showing the record chart and also the tape reader and the punched tape which controls the machine. The recording is made on a strip chart. This chart is advanced by a small step for each measurement. The measurement is carried out by the stylus cluster traveling across the chart. For each measurement the machine makes a nominal mark (small black dots arranged in a straight line) and an actual mark. The distance between the nominal and the actual mark shows the error. In this particular installation each line on the chart is equivalent to .0001". In some of our machines each dot is numbered by an automatic stamping mechanism. In others, this is not found necessary.

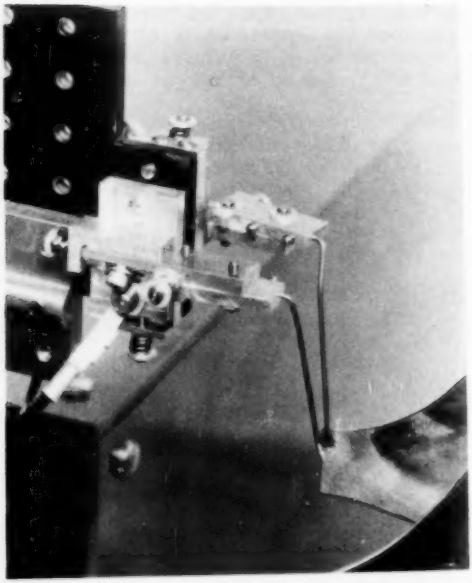


Fig. 3

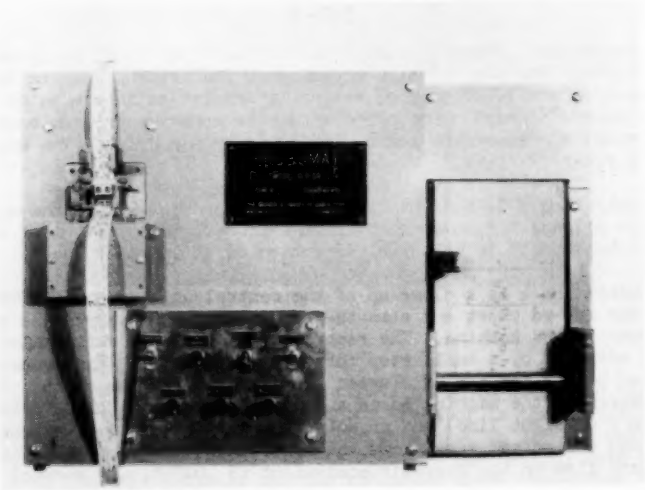


Fig. 4

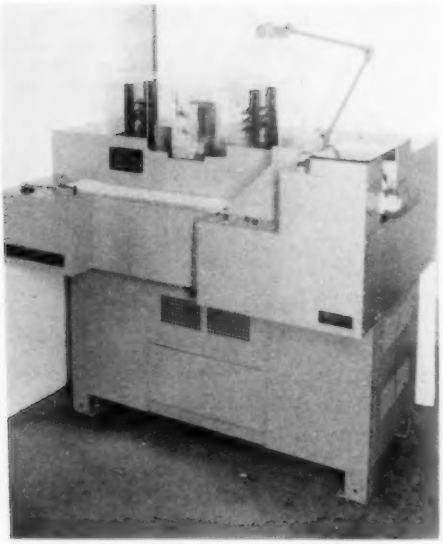


Fig. 5

Illustration 5 shows a front-view of AUTOMATIC AIRFOIL PROBOGRAPH MP-15 which is used primarily for the inspection of large quantities of turbine blades, compressor blades, etc. AIRFOIL PROBOGRAPH MP-15 is an automatic machine. Measurements are taken at 40 or 50 points simultaneously. The measuring results are automatically computed and a green light signifies acceptance, a red light rejection. The machine contains tolerance units which are adjustable to specified tolerance requirements. If all measurements of a blade are within the set tolerances, the machine accepts the blade by showing a green light. If any deviation is outside tolerance, the machine rejects the blade by showing a red light. In this case it automatically repeats its measurement and records all measuring deviations on the strip chart. No record is made of good blades, however the machine can be set to record all or any blade, if desired. The strip chart roll is stored on the right, the chart extends under the stylus unit, comes down the slanted portion and stretches over the whole length of the machine to permit clear and easy observation. The probe banks mount the forty probe holders which are set by pin masters to the precise nominal contour of the blade. The blade is held in the holding fixture visible between the probe banks.

Illustration 6 shows a close-up of the AIRFOIL PROBOGRAPH MP-15. Strip chart and styli can be seen in the foreground. A number of blades are spread out. The probe banks and holding fixture are visible towards the center right. In operation the operator presses a start button; the

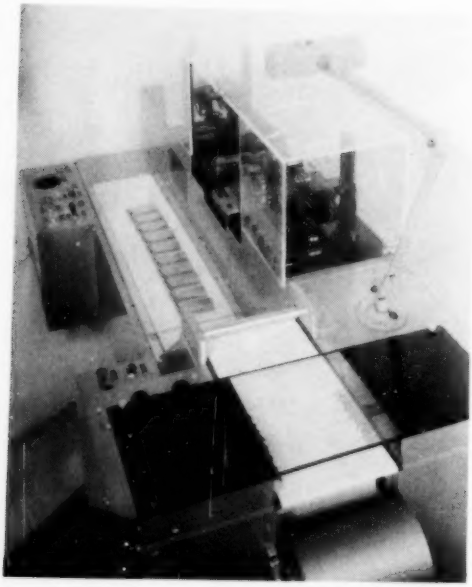


Fig. 6

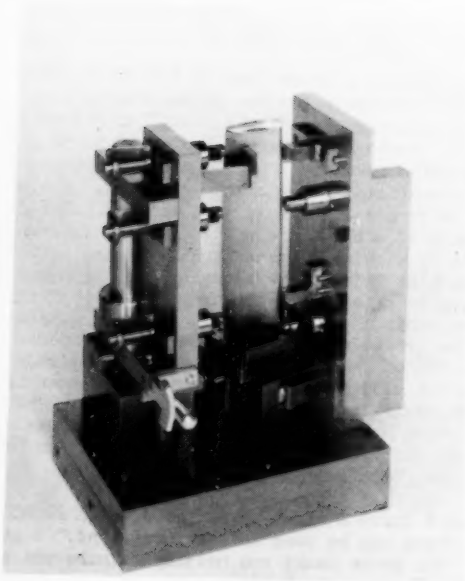


Fig. 7

probe banks approach, first fast, then slow; the probes make contact and take readings of the blade which are recorded by the styli moving over the recording paper at a magnified ratio (40:1, 80:1, etc.).

Illustration 7 shows a blade holding fixture for AIRFOIL PROBOGRAPH MP-15. These fixtures are designed and made for mass inspection to facilitate precise holding and quick loading and unloading. A compressor blade is shown in the fixture.

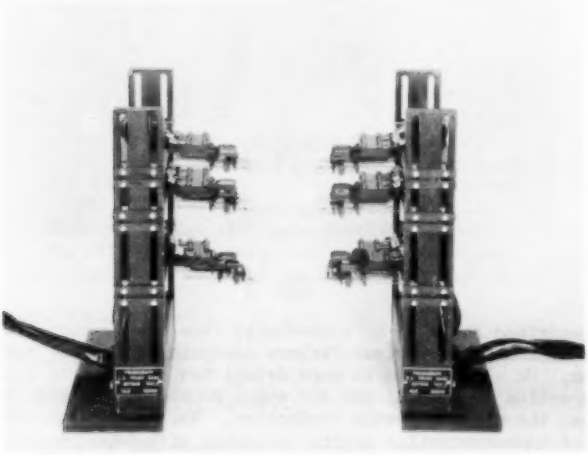


Fig. 8

Illustration 8: The measurements are carried out by probes mounted in banks. This illustration shows a set of probe banks for AIRFOIL PROBOGRAPH MP-15. These probe banks are standard accessories. The probes are held in holders which are mounted to platforms, and these in turn are mounted on the probe bank stands. The probes are wired into a cable with a multiple prong plug. To set up a probe bank is largely a job of assembling standard parts. Probe bank components can be used over.

Illustration 9 shows a set of PROBOGRAPH pin masters. Pin masters are used to set up (master) multi-probe machines. The probing points are represented by adjustable pins. The tips of these pins are ground to the proper contour angle. Pin masters have proven extremely useful. Master blades, of course, could be used to set up a PROBOGRAPH, but master blades are often not perfect and too expensive. They are not adjustable. On the pin master, not only the offset dimension but also the exact station dimensions are fixed. This is very important in interpretation and automatic assorting.

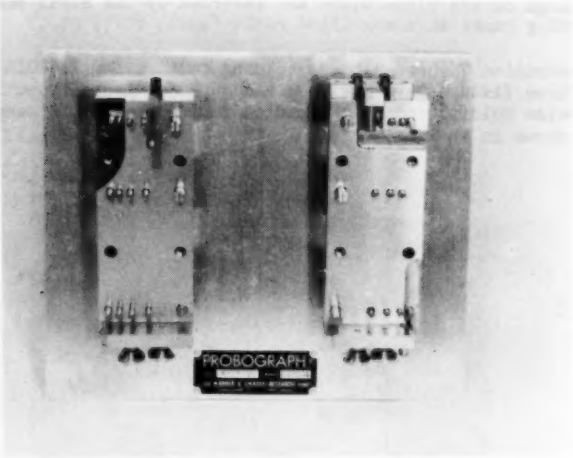


Fig. 2

We described above in an abbreviated form the development, the underlying fundamentals and various factors governing automatic gaging of complex parts. We also showed in more detail two examples of existing automatic inspection equipment, one for small production or even single applications, the other for mass production. We further discussed the close relation of tape-controlled gaging technique with tape-controlled production technique by means of the single-probe type PROBOGRAPH or PROBOMAT as applied to small production. It is interesting to observe that no such close connection exists between mass production on the one hand, and mass inspection by multi-probe technique on the other hand.

This brings up a fundamental difference between small production methods and mass production methods. When speaking of small and mass production we realize that the borderline is not easily drawn. However, there is one basic difference: small production means frequent changes and flexibility of equipment; large production means practically no changes, and durability and efficiency of equipment. Tape control is one of the greatest steps forward in attaining automatic equipment which at the same time is flexible. The equivalent part in mass production is played by transfer machinery. As mentioned before, the inspection problem of tape-controlled machine production can be solved by tape-controlled gaging. The inspection problem of mass production can be solved by the use of tooled automatic computing and recording gaging machines. It will be noted that the gaging methods described here are all applied to the part after it leaves the production machine. No in-process gaging was described. The reason for this lies in the fact that we are dealing here with complex parts only for which inspection in process with feedback correction is not likely to be practical in the near future. On the other hand, automation of gaging of complex parts is a necessity in order to keep pace with automatic production methods. It looked for a while as though automation of production would completely outdistance automation of inspection. We believe we are on the road not only to hold our own but perhaps even to gain the lead.

CONSUMER-SUPPLIER QUALITY PROBLEMS AND RELATIONS

John L. King
Ford Motor Company

The Automotive Division, American Society of Quality Control has a group of ten Task Force Sub-Committees reviewing the use of scientific methods of Quality Control in various phases of the Automative Industry. One sub-committee, of which I am the chairman, is investigating consumer-supplier quality problems and relations.

The stated goals of this task force group are:

1. An Automotive Industry survey to ascertain:

   a. What the major automotive suppliers consider the paramount criticisms of the consumers in quality matters.
   b. What the automotive companies consider the paramount criticisms of the suppliers in quality matters.
   c. What the consumers and suppliers consider the most workable relations they have developed with individual firms, citing the significant things about such relations that cause them to stand out as different.

2. Industry standardization, if practical, on:

   a. Sampling plans.
   b. Nomenclature.
   c. Acceptable quality levels.

3. Possibility of the consumer publishing his inspection method to his supplier at the time of bid, or thereafter, including:

   a. Sampling plan.
   b. Characteristics covered by the sampling plan.
   c. Classification of the characteristics.
   d. Acceptable quality level for each characteristic.

4. Standardization, if practical, of "Initial Samples" and "First Production Shipment" inspection and acceptance practices.

5. Formalized and standardized inspection and quality agreements, certification agreements, etc.

As you can see, the group has outlined some ambitious goals to be achieved and, as you may suspect, these goals have not (at the time of the writing of this paper - 3/8/57) all been achieved. Significant steps toward these goals have been made however, so I would like to present to you what in effect is a progress report of the achievements to date.

Two questionnaires have been prepared and dispersed. One was directed to a random sample of automotive parts suppliers and the other to a random sample of receiving locations in automotive plants. Adequate returns from the later have not as yet been received as the questionnaires have only recently been mailed. A good response from the parts suppliers have been received, however.

Questionnaires were mailed to six hundred parts suppliers. To date, 132 have responded and responses are still trickling in.

It was encouraging to the committee to discover that 94 percent of the responses indicated that the committee's goals were thought to be worthwhile and that our committee was the proper organization to be working toward them. Of course we do not know what opinions the 450 plus suppliers that did not answer the questionnaire have. Maybe 94 percent of them think the opposite. This is probably not the case, however, for if it were they undoubtedly would have returned the questionnaire with that indication.

You may be interested in knowing the size of the companies that responded, which is presumably also the size of those contacted. 34 percent had fewer than 200 employees, 39 percent had from 200 to 700, 16 percent had 700 to 2,000, 9 percent had 2,000 to 5,000 and only 2 percent had over 5,000.

The suppliers were asked if they employed a group of people whose primary function was to control the quality of parts produced by others. 78 percent answered yes and 21 percent answered no. Of those answering no, the majority (67 percent), as might be expected, were in the smallest size catagory, (200 or fewer employees). 26 percent were in the 200 to 700 catagory and only 7 percent above 700. The lack of quality control personnel being in the smaller shops would be expected because management is in closer contact with the workers and the esprit de corps is apt to be more prevalent than in the larger more impersonal plants.

A scatter diagram of the size versus ratio of quality control people to production people indicated some positive correlation (the larger the company the larger the ratio). This correlation was slight, however, and even some of the largest companies had ratios of 1 to 75 or less. And some of the smallest had ratios as large as 1 to 6.

There was a good correlation, however, between the ratio of quality control people and the type of product produced. As one would expect, the precision parts manufacturers had significantly higher ratios than did manufacturers of less precise parts.

The percentage of responses indicating that Statistical Quality Control was employed as a management tool was about evenly split (52 percent yeses and 48 percent noes). This is quite encouraging except that one again tends to wonder about those shops that did not respond. Maybe the reason they didn't was because they did not have a quality control analyst to which to give the job of filling out the questionnaire, or maybe they were so busy struggling with less effective control methods that they didn't have time to respond. Of the answers from plants that manufacture predominantly precision parts, however, over two-thirds indicated that Statistical Quality Control was used.

Of the suppliers that responded, 44 percent indicated that from 20 to 50 percent of their total production goes to the Auto Industry. 16 percent indicated that less than 20 percent, and 40 percent indicated that more than half of their production goes to the Auto Industry. All but one percent have two or more customers in the Industry.

Regarding the cause of quality problems in the suppliers plants, it seems that the biggest contributing factor is the tendency for machine operators and/or their supervisors toward a relaxing of quality consciousness. This indicates that frequent or continuous quality campaigns are necessary to maintain satisfactory quality. However, incapability of machines is sometimes the cause for quality problems as is the nondetection of poor quality parts when they are produced.

In reviewing answers to questions that dealt with the suppliers relations with his customers, it was discovered that inconsistent customer practices and lack of information regarding what is expected of them in the way of quality seemed to be the major areas causing strained relations.

Thirty seven percent indicated their customers were inconsistent in their receiving inspection practices while 22 percent indicated they did not know if their customers were consistent or not. Over half indicated inconsistency between customers while 65 percent felt that it would be desirable to have consistency between them.

It was interesting to discover that 84 percent indicated they believed their customers quality specifications were "generally" reasonable, 3 percent believed them "always" reasonable and the remaining 13 percent indicated them "sometimes" reasonable.

When asked the question whether they had sufficient information regarding what their customers expected of them in the way of quality, only 14 percent answered "always," 70 percent answered "generally" and 14 percent answered "sometimes." Another way of stating this is that 86 percent indicated that they do not always have sufficient information regarding what is expected of them. This percentage is quite high and it appears this is one of the major areas where improvement is needed.

The suppliers would also like this area improved for 91 percent indicated they would like their customers to publish their inspection method including the amount of inspection and the ranking of relative importance of dimensions at the time bids are asked for or thereafter. Several crossed out "or thereafter" indicating they would definitely like to see it done at the time bids are asked for.

We received the indication that most of the time the suppliers have sufficient opportunity to discuss their quality problems with their customers and when discussed they usually arrive at reasonable understandings. It would naturally be desirable for them to always have sufficient opportunities to discuss their problems with their customers when they have problems. It would be even more desirable, however, for both parties if they did not have any problems.

The committee attempted to determine how the suppliers feel regarding Quality Certification Agreements. In the first place there were 29 percent that had entered into such agreements with their customers. Of the ones that were parties to the agreements, only three-fourths felt that they were of definite benefit to them. Approximately one-fourth felt that they had been pressured into making the agreements because they would primarily benefit their customer. Of those not having certification agreements, about half (46 percent) believed that such agreements could be beneficial to them and about two-thirds (62 percent) felt

that such agreements could be beneficial to their customers. It appears that Certification is a field that can stand improvement and expansion and may well bear further investigation by both supplier and consumer.

It seems that the Automotive Industries' practices relating to inspection and acceptance of "Initial Samples" and "First Production Shipments" do not present any major problems to the Suppliers although there is a slight problem there. 17 percent indicated that existing practices were unsatisfactory and that it would be beneficial if an industry wide standardization of such practices were affected.

Fifty four percent of those answering indicated that their customer relations were getting better, 37 percent indicated no change and 9 percent indicated their relations were getting worse. The Committee tried to discover if there were any answers given by the 9 percent that were significantly different from the remainder. Actually, there were not enough in the catagory of "getting worse" to give any significant answers. There did seem to be a tendency toward real differences in some answers to some of the questions, however, and if our sample had been larger we may have discovered significance in the following. It seemed that a larger percentage of their total production goes to the Auto Industry, they feel they have less opportunity to discuss their quality problems with their customers and that they are less apt to arrive at reasonable understandings when they do discuss them, their quality problems are more often caused by a general relaxing of quality consciousness on the part of machine operators and/or their supervisors, and lastly they feel their customers are less consistent in their receiving inspection practices. In analyzing the written comments (all but one contributed additional comments to the questionnaire, the over-all average was one out of four.), it seemed that the major cause for the deteriorating trend in their customer relations was that they felt the quality requirements were continually increasing and that their customers were not willing to pay for the resulting added costs.

In summarizing the answers to the specific questions and drawing conclusions to the written comments, the committee feels the following statements are true. In general:

1. Suppliers do not have enough information about what is expected of them in the way of quality.
2. A small percentage of the specifications are considered unreasonable.
3. There is room for expansion for Statistical Quality Control.
4. Machines are more capable than men in the job of producing quality parts.
5. There are inconsistent practices in and between receiving inspection areas and it would be advantageous if these were eliminated.
6. There is room for improvement and expansion in Quality Level Certification practices.
7. "Initial Samples" and "First Production Shipment" inspection and acceptance practices are fair but can stand some improvement and standardization.
8. The customers are becoming more and more demanding. Competition is becoming keener. Quality requirements are steadily increasing.

9. And lastly, the key to good customer relations definitely seems to be good communications. It seems that if the Suppliers maintain close contact with his customers' Engineering and Quality Control personnel he generally avoids having trouble or if he does have any, he soon gets out of it.

In conclusion, the Sub-Committee considering consumer-supplier quality problems and relations plans to do further analysis of the results of this questionnaire and the results of the questionnaire sent to the auto manufacturers (when adequate results are received). We plan to have additional information regarding these questionnaires and the Committee's recommendations available for presentation at the Quality Control Convention in Detroit, May 22-24, 1957.

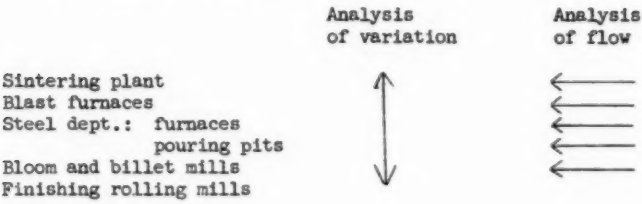
REPORTING AND ANALYZING PRODUCTION IN A SWEDISH STEEL PLANT.

C. Jan Yngstrom
Domnarfvet Iron & Steel Works.

In order to get a rationalized production and a stabilized quality statistical methods and statistical thinking have proved to be useful tools. It has been possible to use methods worked out by theorists. In applying the theories in practice the U.S.A. have been pioneers and the A.S.Q.C. have been active contributors. It is a great pleasure to state the fact that the American results have been widely published and they have been leading the development. The intention here is now to show in what ways we have been proceding in a Swedish steel plant to stimulate statistical thinking.

The examples here are taken from the Domnarfvet Iron & Steel Works, owned by the Stora Kopparberg Company, Sweden. The production 1956 was 550.000 tons of ingots produced in four basic Bessemer converters, a kaldofurnace and five electro-furnaces. All of them are charged with heats of about only 30 tons. The number of heats was therefore comparatively great, totalling some 22.000. The products are hot rolled plate, rails and beams, sections, rods and strips.

A steel plant is - as we all know - a continuous chain of production and the variations must be analysed in each section. Each department must work on whatever the previous department hands down. That means that the flow must be canalized and easily analysed.



The Human Factor

It has been found that most mistakes are made owing to insufficient instructions or to people being out of balance. It happened in the sinter plant that the additions of limestone were not all right. The additions are made by four particular men called mixers, one on each turn. The variations occurred specially on the night turn. The result for two months is shown in the following table.

Men in charge	No. of observeations		Mean deviation from lime ratio wanted
	-	+	
Mixer No. 1	54	67	0.08
2	88	16	0.07
3	74	39	0.07
4	82	29	0.11

Mixer 1 had been operating quite well. Mixers 2-4 had been operating too cautiously and they were given new instructions. Moreover, Mixer 4 had too big a mean deviation, as compared with the three others. He happened to be an elderly man, who had just had some kind of breakdown and he was at that time not fit for the night turn.

The example shows how the human factor can influence production. It is therefore necessary that the reports all over the plant clearly show who is responsible for the different operations.

Technical Relations can be Overshadowed

A technical process is very often disturbed by secondary factors. These factors can be of such a magnitude that it is impossible to get a clear view of the technical relations. An example is shown in figures 1. and 2. and it is taken from the blast furnaces. The intention was to see if there was any relation between the manganese and sulphur content in the pig iron produced. Figure 1. contains all observations from a certain period. The figure is a bit confusing. Figure 2 shows only the values representing 24 hours without any special delay and with silicon and basicity on normal level.

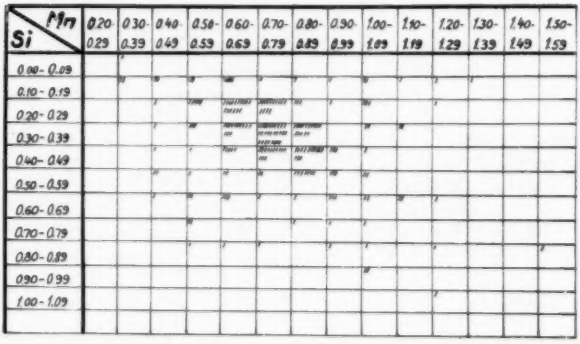


Figure 1.

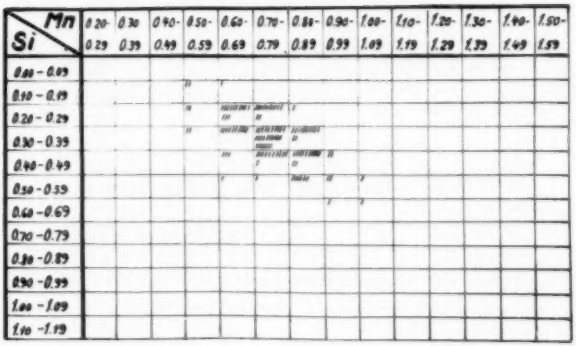
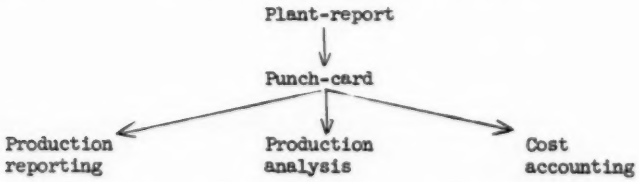


Figure 2.

In this case only two factors were studied. Normally, there are several factors such as consumption of raw material and fuel. It seems advisable not to use all figures available but just the observations where known disturbances have been eliminated. If one wants to study the influence of disturbances or delays on total production, this will have to be done separately.

Reporting Routines

When the production is uniform, as it is in the sinter plant and in the blast furnaces, the reporting routines can be quite simple. But when production is split up in various grades and products, the reporting is a bit more complicated. All figures wanted can, of course, be extracted manually. Since the advent of the punch-card system the figures are more easily made available for both reporting and analyzing.



It is noticeable that the system described here was developed and put into operation in Wheeling Steel Co. by Frank G. Norris almost at the same time as in Domnarfvet. - This system has proved to be very good in following up flow-channels, as a controlling instrument as well as for studying technical relations.

Complicated Reports are not wanted

If the reports are too complicated, they very soon become inattractive, and will not be studied with sufficient interest. Figure 3.

shows a very simple example of a report from the Plate Mill, where cracks, scabs and other defects are indicated. One copy goes straight to the metallurgical section and the other goes to the inspection department.

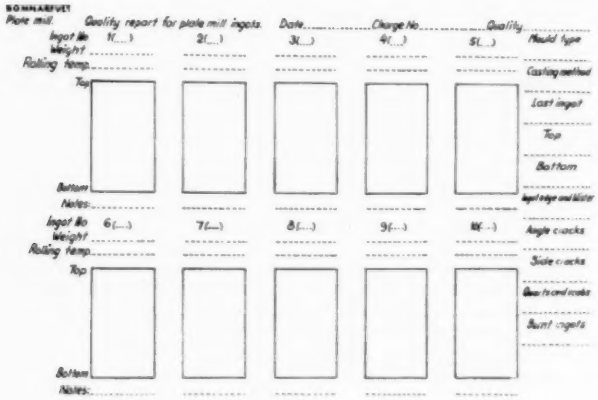


Figure 3.

Figure 4 shows the yield in one of the rolling mills and it gives a true picture of the influence of the defects on the yield. It also shows the yield-level for defectless material.

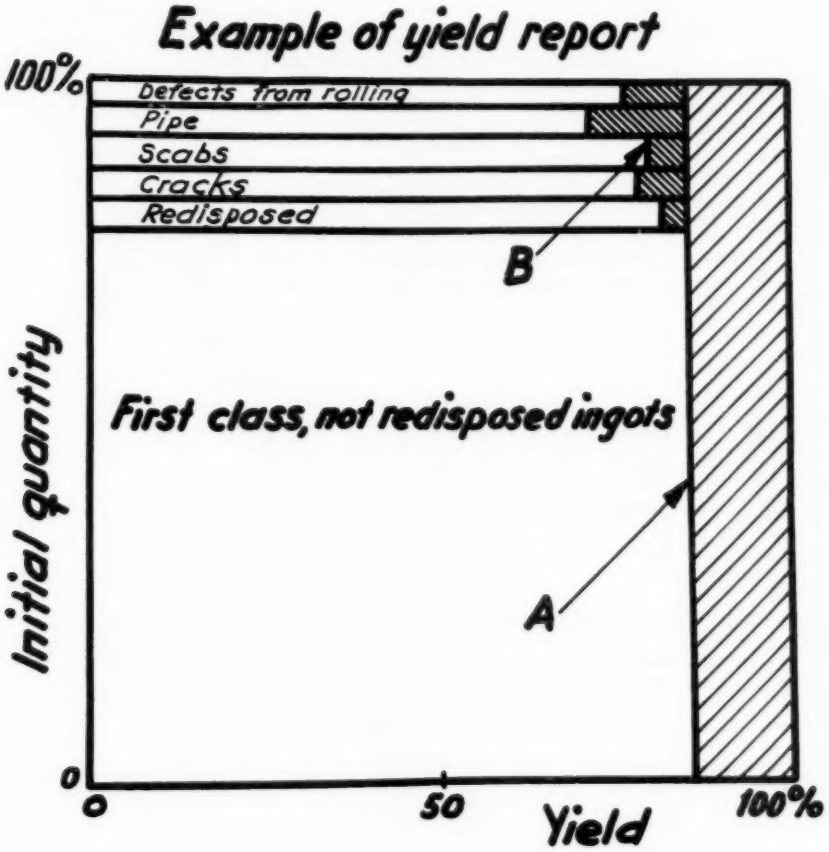


Figure 4.

Centralized or Decentralized Reporting

Previously, there was a trend to centralize the assembling of the plant reports, the idea being to reduce hands, but, owing to the impossibility of keeping sufficient contact with the plant-personnel, the control proved to be inadequate and a considerable delay in completing the reports was observed.

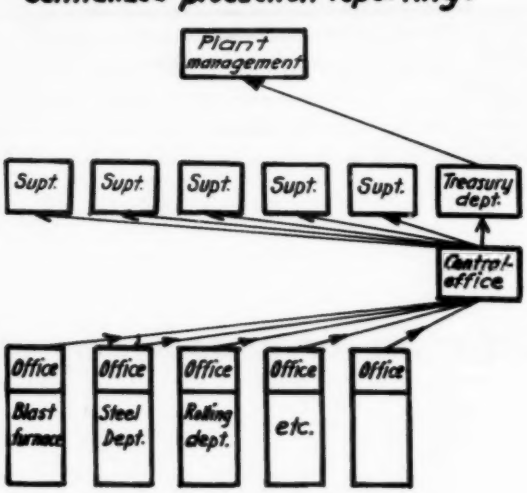


Figure 5.

e development has definitely gone towards a decentralized system
6.) The reports are made by the different department offices -
erintendents being reponsible for them. The reports passing
ards are also put together by the department.

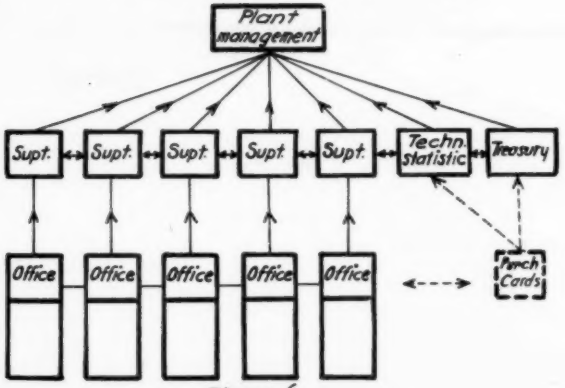
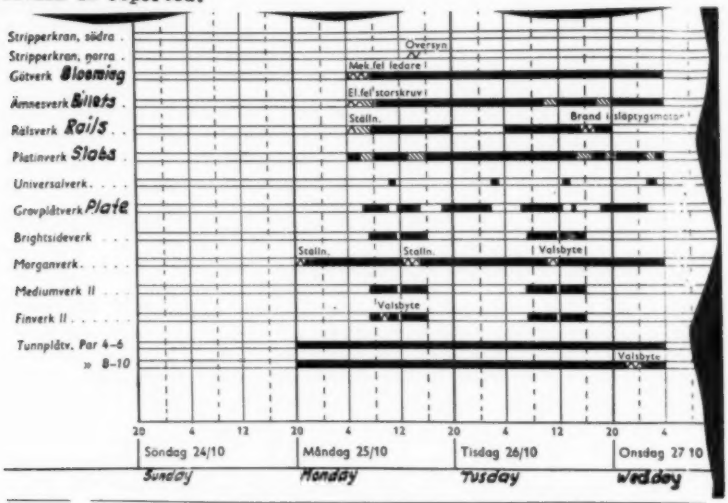


Figure 6.

In order to get a similar reporting system all over the plant it is necessary to co-ordinate the different routines for which the technical statistics is responsible. All new routines or deviations from old ones have to be checked to fit into the whole system.

Reporting to Management

The Management must be kept informed about the actual situation. For each unit, operating time and delays are reported graphically. The production reached is put in relation to the scheduled output, and the deviation is reported.



	Söndag 26/12	Måndag 27/12	Tisdag 28/12	Onsdag 29/12	Torsdag 30/12	Fredag 31/12	Lördag 1/1
Blästerhyttorna Blast furnaces	Sunday	Monday	Tuesday	Wed.day	Thursday	Friday	Sat.day
Erhållet	1.182	1.312	1.185				
∴ kapacitet	− 182	− 312	− 185				
Cum. erhållet		2.494	3.679				
∴ cum. kapacitet		− 494	− 679				
Thomasgöt Basic bessemer							
Erhållet	1.311	890	809				
∴ planerat	− 211	− 110	− 29				
Cum. erhållet		2.201	3.010				
∴ cum. planerat		− 321	− 350				
Elektrogöt Electro furnaces							
Erhållet		591	722				
∴ planerat		− 159	− 218				
Cum. erhållet			1.313				
∴ cum. planerat			− 377				
Upplagda göt Stored ingots							
Per dygn	1.228	181	0				
Cum.		1.409	1.409				
Nedsatt kalla göt							

Figure 7.

These reports are made for every 24 hours and the sheet includes a week's result. They are very simple but have none the less proved to be of great help to the Management. The weekly and monthly reports are traditional, and, unfortunately, they tend to become much too extensive, but efforts are being made to reduce their number as much as possible.

Technical Statistics

What, in the plant, should be subject to special studies, and who shall do it? The General Superintendent has to decide what is the most urgent problem, and the study should be made in close contact with the superintendents involved. A small department called Technical Statistics has been set up.

An interesting problem which we have given a lot of thought is the relation between the chemical analyses of steel and the physical properties of the finished product. Earlier practice has been based on a carbon value arrived at through theoretical and empirical investigations. It has been:

$$C_v = \left[1 + 0.5\ (C-0.20)\right] \cdot C + 0.15\ Si + \left[0.125 + 0.25\ (C-0.20)\right] Mn$$

+ 1.25 P and the carbon value gives the physical strength according to the formula:

$$\sigma_s = k_1 + k_2 \cdot C_v$$

where k_1 and k_2 are based on the values from the plant, and they depend upon the dimensions of the finished product etc.

Instead of starting from the theoretical carbon value the multiple correlation between σ_s and the analyses for C, Si, Mn and P has been calculated. The same thing has been described in Industrial Quality Control No. 4, 1956, by J.F. Occasione, and it is interesting to see how similar the problem has been. For 8 mms concrete reinforcement bars and with the following mean of the analyses

C = 0.26 Si = 0.27 Mn = 0.92 P = 0.45

the equation was:

$$\sigma_s = 18.20 + 50.67\ C + 9.04\ Si + 8.84\ Mn + 54.7\ P$$

or, generally written:

$$\sigma_s = k_1 + k_2 C + k_3\ Si + k_4 Mn + k_5 P$$

If this equation is transformed to make a comparison with the formula earlier used, we arrive at the following:

$$C_v = 1.09\ C + 0.19\ Si + 0.19\ Mn + 1.18\ P$$

and the C_v for this grade is according to the earlier formula:

$$C_v = 1.03\ C + 0.15\ Si + 0.14\ Mn + 1.25\ P$$

The comparison shows that the earlier formula is not quite all right. The new values gives the situation for the produced steel and the rolling mill facilities at that time.

The next question was if the constants for C, Si, Mn and P (k_2 + k_3 + k_4 + k_5) would change, if the rounds were not 8 mms but 16 or 25 mms. They proved to be about the same, except the first constant factor, which changed. For 16 mms it was 12.0 and for 25 mms 9.7. The factors k_2 , k_3, k_4 and k_5 can then be calculated as averages for all sizes of concrete bars of that grade. It is now practical to make a new carbon value

$$C_{v_1} = \overline{k}_2 C + \overline{k}_3 Si + \overline{k}_4 Mn + \overline{k}_5 P$$

and the relation between the chemical analyses and σ_s is

$$\sigma_s = D + C_{v_1}$$

where $D = k_1$ is a measure of dimension.

It is possible to make a diagram of the relation between C_{v_1} and σ_s, where the spread of the values can be taken into consideration.

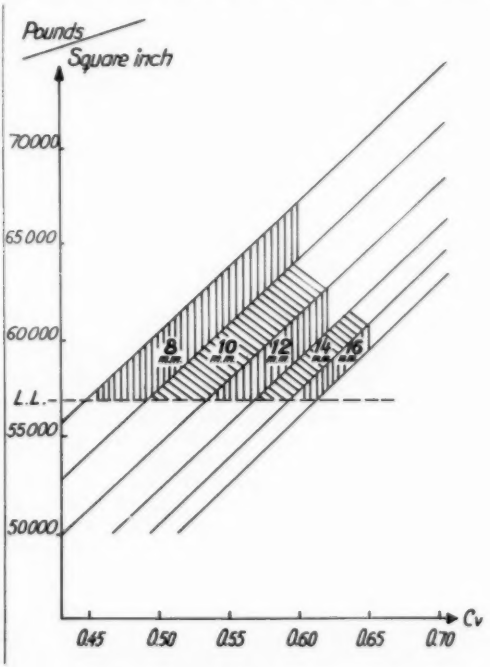


Figure 8.

When the heat is in the soaking pit and the rolling mills get the carbon values from the lab, it is possible to decide whether the heat can be rolled to the intended size or if it should be stored as billet and rolled on another dimension.

What then are the objections with regard to the exactitude of the results? The testing of σ_s has a spread of about ± 0.5 kg/mm2. Further, the chemical analysis of a single rolled bar deviates often quite a lot from the analysis of the heat. Then the equations were counted from the analyses of tested bars and not from the analysis of heat. Finally, the rolling temperature in the finishing mill must be constant since it has great influence on the physical strength. Development, however, goes forward to a stabilization of production. The equations showing the relation between chemical analyses and in this case physical strength will also serve as an instrument controlling production from molten steel to finished product. The carbon value can even be useful in connection with the work of minimizing the number of grades in the plant.

This study has, like most of them, not ended up in anything definite. It has, however, shown that the equations and particularly the carbon value are useful tools in controlling the production.

Closing Remarks

Statistical thinking is nothing new, but it has now been systematized and it is accepted not only by Management but also by technicians in general. The cooperation has been good, mainly because the work has been carried out quite openly and not in any form of secrecy. In Sweden, the Swedish Iron Masters Association have arranged a statistical bureau to keep in contact with the steel plants. The question how to train technicians for statistical work is actualized, and the way by group-information, as it has been practised here in the U.S.A., seems to be the best. The A.S.Q.C. has in a marvellous way promoted this development.

STANDARDIZATION AND QUALITY CONTROL

John J. Dunn
Office of the Assistant Secretary of Defense (Supply & Logistics)

It is indeed a pleasure for me to be afforded the opportunity to participate in the eleventh annual convention of the American Society for Quality Control. I propose to take advantage of this opportunity to examine with you some features of inspection and quality control in which I have a very direct concern.

My interest in quality control stems from my responsibility for the Defense Standardization Program in which the technical requirements for quality control are of great importance. The Defense Standardization Program undertakes to reduce the variety of goods and services required for logistic support of our operating military forces in four basic ways:

1. By reducing the variety of items presently in the supply systems: through the elimination of duplicates, near-duplicates, and of unneeded and obsolete items. This process we refer to as "simplification".

2. By the development of engineering standards for items and their components which will limit selection in the design of new items.

3. By standardizing practices and procedures associated with the design, development, production, shipment and use of items. In this connection, standards are developed for such things as drafting room practices, packaging, quality control, etc., when these practices are applicable to many items.

4. By the establishment of uniform requirements in the specifications used in procurement and the standardization of practices and procedures peculiar to individual items.

It is, of course, in the latter two areas in which my interest in quality control lies -- standards for established practices affecting a great number of items, and uniform requirements for the quality assurance of individual items.

Before identifying the nature of my concern with quality control, I should like to have the record show that I am quite aware that, in many respects, the art of quality control, as applied to the purchaser rather than the producer, has not attained the degree of stability which offers the essential ground for the development of standard practices. Yet, I suggest that the degree of stability is perhaps much greater than is recognized by the profession itself. In your natural concern to develop and perfect the techniques which distinguish this profession, you have perhaps done yourselves a disservice by having given insufficient attention to abstracting from your experience those policies and principles underlying your work which would further promote and solidify the gains already made in this field.

My interest in quality control, therefore, is to translate to the maximum extent an art into a science, so that those functions of management for which you develop your techniques are enabled to employ them most effectively to improve the activity in which they are employed

and enhance the profession which provided them.

If my remarks are to be understood, it is essential for me to make my position clear in the context of my own responsibilities and interests in this area. Obviously there is a difference of degree and, in some measure, of kind, in the interest which people involved in the production of materiel have, as distinguished from the interest which must exist on the part of the Department of Defense as a consumer of this materiel. In this connection, therefore, my concern stems from the necessity for providing military specifications which set forth clearly and accurately our requirements, utilizing where we can, the tools of statistics and quality control engineering.

Let me further clarify the position of the Department of Defense in this matter. Obviously, industry is concerned with quality control as a technique for the in-process phases of production of an item. Its concern properly is fabrication of an item from the raw material through the finished product. The major purpose of quality control in this effort is to determine the factors which will result in a suitable material capable of being produced and sold. For this reason it is essential to the producer that every process variable be studied and its effects anticipated. Quality control in this sense is directed to the avoidance of defective materials and to the production of items complying with the requirements of the purchaser or of the designer. On the other hand, the interest of the consumer must obviously be to build quality assurance requirements into his purchase specifications which will, while insuring a suitable product, still minimize the cost and impact of inspection and testing. To the extent this is feasible, an item containing the optimum amount of utility is made available at least cost.

It should be obvious that it is to the mutual interest of the vendor and purchaser to have specifications that are rational and intelligently stated. There will then be a minimum of misunderstanding and a minimum of waste of time, effort and resources involved in the production of all items so specified. This is the point which we feel is of mutual concern to you who are present here today and to the Department of Defense. That is, in the evolution of control techniques to avoid producing defective materials, tools have been developed which minimize the cost of production of each item by applying to the total number of items produced, the intelligence gathered from previous experience in production. It is my belief that the unfortunate situation of the specification developer today is that the same intelligent application of effort has not been given to the preparation of quality assurance requirements of purchase specifications.

Statistical techniques, control limits, process variables, and other such criteria have been useful in the development of the proper measure of control over material in the production end. It should be apparent that it is to the interest of the Government to have a similar type of scientific control and approach to the development of specifications. This has not been given sufficient recognition in the past. More attention must be paid in the future to the application of the same types of resources which have been applied to the control of production to assure that specifications state in the same definitive terms the minimum requirements which need to be stated. This will assure sampling plans being on a rational basis and that classifications of defects and acceptable quality levels are defined within limits developed through

the media applied in the past to production quality control techniques.

With the development of the appropriate mathematical and statistical and related techniques in the development of acceptable quality levels, classifications of defects and rational sampling plans, it is anticipated, in lieu of the haphazard and conglomerate nature of present sampling plans or techniques for selection of such plans or for the delineation of limits of acceptability, that patterns, principles, and uniformity will emerge. This will place quality assurance provisions on a rational basis. The same intelligence can then be applied to the methods of developing quality assurance techniques and provisions as has been characteristic of the gradual improvement of production quality control techniques.

This, we feel, is a matter of joint interest since it is obvious to us that any irrationality which creeps into our specifications affects you directly as the contractors and vendors who must furnish materials in compliance with Government specifications. To the extent that we can make our approach as rational as the approach to quality control in production, we will not only improve our relations with industry but we will reduce the costs of failure to communicate intelligibly and adequately with the people who are expected to serve the Government's needs. It is not easy to face a situation in which the will to correct any deficiencies is as great as it is in the Department of Defense today, and yet to be frustrated because the techniques and tools which should be available to be picked up are not recognized. Our awareness has developed in this area but the means to correct the situation are unfortunately so hidden or else so dedicated to only one aspect of the problem that all our efforts in this direction are not on a planned and rational basis but are haphazard and empirical. This results in such a variety of requirements being placed on the industry, often on the one industry by a number of requiring activities, as to constitute a very real hazard to proper working relationships between the Government and Industry.

I am therefore addressing myself today to the task of bringing to you the questions which exist in our minds. I hope that development of these problems will stimulate added thought on these questions. In that way, any irrationality which may exist in the military specification system will be removed. We recognize military procurement as one of the largest factors in industrial production today, and stability in this area is an urgent necessity.

I should like to illustrate the theme of my talk by examining two features of this circumstance -- one, what guidance has the quality control engineer provided the designer in the selection of those techniques of quality control which will give the designer the needed assurance that his item, if it meets these requirements, will in fact perform as designed? -- and two, has the quality control engineer, in the absence of this guidance, unconsciously assumed some share of responsibility for the performance of the item which the designer cannot properly share with anyone else?

In considering these questions, let us look at AQL's and CD's. An analysis of military specifications today will disclose that a wide variety of quality control requirements are employed. These will vary from a simple requirement to test "2 out of each 100" to the most complex array of visual and dimensional examination procedures; of

acceptable quality levels with accompanying sampling plans; of classifications of defects for major, minor and critical features of parts, materials and end items and the like.

While the variety of techniques selected to assure quality are not subject to criticism, since this may reflect only the ingenuity of the profession, what is distressing is that there is no discernible pattern among these many techniques which would assist in establishing criteria for their most effective employment. Are these requirements conditioned by the nature of the item, by the nature of the industry, by the application for the item, by the method of fabrication, or by other constants? If so, what is the nature and degree of influence which these factors exercise in properly selecting a sampling plan, an AQL, or a classification of defects?

These questions must be answered if management is to be enabled to uniformly and effectively employ the fruits of your labors. While your energies are quite properly devoted to the development, validation and refinement of your techniques, criteria for their employment in the area of quality assurance as well as in the area of quality control should hold equal priority in your research.

Again, looking at classifications of defects as a tool of quality assurance, we find that the nature and content of these documents show extreme variation. One theory concerning these documents is that they are designed to provide the specifics of "good workmanship". By providing definition to this technical requirement, the area of acceptance judgment required of the inspector is accordingly circumscribed. This is good, of course, since it minimizes differences between the contractor and the customer and increases the chances for equal treatment of all vendors.

Yet you will find in many of these documents quality control requirements which go far beyond those associated with workmanship. There are requirements with respect to dimensions, materials, and other technical requirements. The requirements for these features of an item are normally called out in Section 3 of the specification.

With respect to those considerations which have a direct bearing on the pricing of a product offered to the Government, there are two extreme viewpoints on specifying AQL's, sampling plans, and CD's in specifications. One viewpoint is to exclude them completely from specifications. These provisions are then made available only to the Government inspector who may need them to check the products for acceptance or rejection. The supplier and the Government inspector are not limited in their relations by rigid and detailed quality assurance provisions. The supplier will have a free hand in establishing whatever inspection and quality control system he considers necessary. This procedure would avoid an interpretation of the specification by the supplier that a rigid application of fixed inspection procedures by the supplier is required.

The second viewpoint is to prepare complete and detailed quality assurance provisions for specifications. Aside from the fact that it is the policy of the Department of Defense to prepare complete and detailed provisions when the necessary data is available and pertinent, there are definite economic advantages to this procedure. The supplier will have guidance as to the inspection procedures to which his product

may be subjected; the supplier will have an equitable basis for estimating costs for bidding on Government contracts; the same inspection standards are used for all suppliers of the same item; and Government - Industry relations are improved when inspection standards are clearly and uniformly defined.

An examination of all the information that has come to our attention in this area has made it abundantly clear that three questions have to be answered:

1. What are the engineering and statistical criteria for specifying an acceptable quality level (AQL) and sampling plan?

2. What are the engineering and statistical criteria for specifying the degree of importance of the quality characteristics of an item for inspection purposes and classifying defects as being minor, major, or critical?

3. What are the economic factors which are involved in specifying an acceptable quality level and sampling plan?

Many rules-of-thumb are used in specifying AQL's. A single standard value may be used for a large class of materials or commodities, irrespective of the nature of the specific item. In other cases, the AQL represents an intelligent estimate on the part of the development engineer or specification writer. For some items there is a wealth of quality history records from past procurements for estimating realistic AQL's. But when this information is not available, the AQL specified in the specification may not be realistic. For instance, the AQL may not be in line with good industry practice or with the needs of the DOD. The information feedback from the supplier or Government inspector to the preparing activity may then indicate the need to modify the specification, often a slow and tedious process. Further, this corrective action doesn't help with respect to the general run of new specifications or revisions of specifications which for the first time introduce these new statistical tools for product inspection.

With respect to statistical sampling plans, specifications contain a variety of requirements for specifying lot size, sample size, and acceptance and rejection numbers for products. The lot size may be arbitrarily fixed at a large or small number, or may be varied to cover a day's production or one process cycle. The sample size, often dependent on the lot size, may also be fixed or varied. It is also necessary to determine whether single, double or sequential sampling will be specified. Acceptance or rejection may also depend on total numbers of defects or may be measured by the number of defective products. Thus, there are many possible combinations from which a particular sampling plan may be chosen. The criteria for determining the choice should arise from the deliberations of a group such as this which has done so much definitive work in the area of production quality control.

The other issues to be considered are the CD's used in attribute inspection, the most common statistically-based inspection procedure used in military specifications today. CD's are classifications, in an assumed order of importance, of those characteristics of products which require examination to determine conformance with the technical requirements of the specifications. How to objectively classify these characteristics? How to determine the degree of detail required of CD's? Is the classification as critical, major and minor sufficient or do we

also need the categories such as minor A, minor B, incidental, and special, now to be found in specifications? Without soundly based criteria for answering these questions, there is confusion in this area of specification requirements.

The question of CD's is bound up with AQL's. There is always the question whether the AQL and related sampling plan specified for a commodity are reasonable. The amount of examination specified in the quality assurance provisions may seem excessive when applied to products with a high degree of uniformity; or the specified examination may be inadequate when applied to products that are not of a highly uniform quality. One interesting case which demonstrates this problem occurred recently with respect to bar soap. Soap is recognized as a large-volume product of extremely uniform quality. In one instance an AQL of 4% was proposed. This would require sampling far beyond the number actually required to determine quality of such a uniform product. For such a product which is very uniform, it was indicated a sample of 10 might be sufficient for all the physical characteristics irrespective of lot size. Therefore on what basis is an AQL decided upon?

Another theory is that classifications of defects may not impose requirements beyond those of the specification. Yet we find many instances where this is done. For example, Military Specification MIL-C-1283A specifies a 5-gallon military gasoline can, known popularly as the Jerry Can.

A close study of the specification and associated production drawings shows that the CD demonstrates as defects forming a basis for rejection under the appropriate AQL of at least 3 characteristics not described outside the CD, although these are reasonable criteria to apply as a test of "good workmanship". In another case the height tolerance varies between the drawing and the CD. On the other hand, MIL-C-13984A specifying a 5-gallon military water can, also the Jerry Can, but for another application, not only raises the same sort of questions, but introduces a conflict in classification between the two Jerry Can specifications by requiring the tolerance on closure seam of the handle to be 1/32" in one can and 1/64" in the other. Here, indeed, is cause for achieving some semblance of order.

Again, are classifications of defects properly a part of the specification and, accordingly, of the contract? Practice within the military varies in this respect. This variation will continue until the nature and function of these technical documents is clearly established, and criteria for their use established.

Let me cite additional examples of diverse quality assurance requirements:

1. MIL-P-35002, Paragraph 4.?.1 - The following statement appears, "The Government reserves the right to classify as a defect any departure from the specification not specifically listed in these tables." This permits the Government inspector to add to or delete from requirements which have been made mandatory upon the manufacturer.

2. MIL-C-13449(Ord), Paragraph 3.1.3 - The table under this paragraph has a column heading "major or minor" defect. No guidance is provided concerning such a classification. In addition, paragraph 3.12 of this specification has, in effect, a second listing of defects which

covers performance testing, Two overlapping listings of defects in the same specification is confusing.

3. MIL-A-13488(Ord), Paragraph 4.2.1.1 - Listing of defects by referencing the paragraphs in Section 3 does not provide the definition necessary. For example, paragraph 3.7.2 states that welding or brazing shall be acceptable as to quality, workmanship, and appearance. Similarly paragraph 3.7.1.

4. MIL-C-13993(Ord), Paragraph 4.1.2 - Reference is made to incidental defects. No guidance is provided as to the difference between minor and incidental defects.

5. MIL-I-6870(ASG), Paragraph 3.4.1.1.1 - 100 percent inspection is required. This Division has no guidance to pass on to the departments as to when to use 100 percent inspection as against less than 100 percent.

6. MIL-STD-152, Section 3, Paragraph 3.2 - The sampling for environmental testing requires that "1 or more items selected at random from the first 10" shall be used. This requires the Government inspector to use his own judgment as to the number of samples to select. Again no guidance is available to this Division which would enable the departments to make more definite statements in specifications on this point.

The previous questions have been discussed from the engineering point of view. Serious consideration must also be given to the economic factors. Supposedly, AQL's, CD's, and statistical sampling plans have received wide recognition as quality control tools because the Government will be assured of obtaining a required quality at a product cost commensurate with its needs. Government inspection costs should be reduced in proportion to the employment of these tools in the production process. Production efficiency is promoted because the supplier is able to gear his operations to these predetermined requirements. These considerations are sufficiently important to justify a careful examination of the economic factors involved in determining quality standards for specifications. The effect of these requirements on the Government and the supplier should also be studied since quality assurance provisions of specifications should be so written as to promote maximum cooperation between the Government and industry.

Questions such as I have suggested in the last few minutes have been repeatedly asked, and the answer has sometimes been "Ask the Quality Control Engineer. He knows." This may be so. But, if this is so, surely he can reduce his knowledge to principles available to all who would know for the designer, the engineer, and technician to select in order to assure the product will perform as designed.

And, if a failure of the handle seem to meet within 1/64" is a basis for rejection, should the designer or the Quality Control Engineer so determine? This is a vital question if an organization is to be able to fix responsibility with certainty.

I submit that the interest of quality control requires that a precise delimitation of responsibility in these areas be established. The future of this profession could be seriously affected by an indefinite continuation of the existing uncertain situation. The

integrity of the role of quality control principles in Government specifications is at stake.

Again, if quality control is to assume some share of responsibility for design, self-interest requires that the extent of its responsibility be defined. Only in this way can it pursue the objectives of its function with steadfast purpose.

The cost of a product to the Government may be materially affected by the quality assurance provisions of the applicable specification. The actual effect will depend on the specific provisions of the specification and the contract. If the supplier is required to inspect the products exactly in accordance with the quality assurance provisions of the specification, he may make an additional charge for this service. If the quality assurance provisions of the specification are considered only as criteria which should insure adequate products and the supplier's own inspection and quality control system, while not directed to the specification, is adequate, there is no additional cost factor and therefore no increase in price of product should be necessary.

The improvements in the techniques for assuring quality and the economy with which these improvements have been attained in the last twenty years is well established. Perhaps now is a good time to consolidate the gains, and establish the benchmark from which future achievement can be measured. Summarizing, then, the simple points I have presented. Identify the nature, function and criteria for application of the quality assurance techniques which have been so well developed. With this knowledge in the hands of management and the other echelons required to employ your techniques, further progress in this field will be accompanied by the understanding and support necessary to the advancement of your profession.

While my remarks may appear critical, the problems which I mention are recognized by the responsible quality control officials of the Department of Defense, and research on these problems is proceeding. This research, however, can certainly benefit from study within the profession generally.

When we examine the economic and engineering and statistical factors concerning the use of quality control techniques in specifications, we feel certain that it should continue to be the policy of the Department of Defense to include detailed and complete quality assurance provisions in specifications. We are also deeply concerned with the need to develop appropriate management policies in this area. I hope that the questions which I have discussed here will command your serious attention so that the criteria needed for such policy will be studied and developed. Without the criteria and the policy, it will not be practicable to promote the most effective use of these advanced tools in the standardization program.

Essential to using these new tools or techniques in standardization is uniformity of practice. It is necessary that the military services apply the same or similar inspection procedures when procuring the same or similar items for the same use. At present, the confusion that exists in writing quality assurance provisions does not permit us to standardize these practices.

In conclusion, it is hoped that the necessary data and information

required to prepare guidance for applying quality control methods and techniques to Government quality assurance requirements and for developing the required policies for their application will soon be available. When that stumbling block has been overcome, we hope to continue to promote flexible but uniform practices, and thereby a higher degree of efficiency in our procurement program.

PROCESS CONTROL IN A JOB-SHOP

Frank Caplan, Jr.
Westinghouse Electric Corporation
Atomic Fuel Department

When statistical quality control techniques were first developed, it was felt by many that these methods applied specifically to large industries which mass-produced great volumes of identical items. A number of years later, medium-sized and even small companies producing some volume of parts, supposedly identical, were making use of statistical techniques to good advantage. Applications in the job-shop, however, up to the present have been few and relatively poorly documented.

The system which I will describe was devised for a development shop where the great bulk of material produced fell in quantities of one to twenty. The parts ranged in size from small gears to large forgings and castings. Types of operations included lathe, drilling, grinding, vertical turret lathe, horizontal boring mill, and sheet metal fabrication as well as hand rework operations. Although there were some reasonably large volume items produced in the shop, these were not covered by this system but a different one was developed for use in those areas.

Since the above-listed operations cover the bulk of machining practice in American job-shops, I believe that we can now throw away the attitude "it's all right for Joe Blow, but it won't work in my shop," and see whether this technique actually can be applied generally.

The first step in establishing the system was to develop and adopt two points of philosophy. The first all-important point was that the operators were not producing parts as such but were producing inspection points. This immediately translated to a common level all operations on all machines and handwork. It didn't matter whether the inspection point was a diameter or the removal of burrs around a drilled hole. In either case, they could be compared directly as representative of whether or not the machine, the man, or the method was capable of doing the work required.

The second important point of philosophy was that we were not in the least bit concerned with punishing people who produced defective inspection points. We wanted to determine only what the cause was and correct it so that defective parts would not be produced in the future. Therefore, a basic part of the system was the requirement that the foreman investigate and report on each out-of-control point.

Each inspector carried a notebook whose pages were made up like Figure 1. In this, he recorded the machine number, how many points he had inspected during the shift, the proper shift to be credited with these points, and the number of deviating points he had found. He then signed each page and turned them in to his foreman. The following morning, the sheets were collected by a representative of the Quality Control Department and tabulated on floor charts similar to Figure 2. These charts were posted on large easels in each department and carried the information for a two-week period. They showed the number of defects found and the number of inspection points examined for each shift, as well as the cummulative defects and inspection points over the period of the chart. These were charted by shift for each machine in the area -- in rework by the type of defect. Since the process average had been

calculated in the previous period, whether or not the indicated results were out-of-control could be determined by reference to a number defective table which sets up control limits for a wide range of sample size and fraction defective. Numbers defective on the high side of the allowable value were circled in red; those below the low limit were indicated in green; and the charts presented a very clear picture of the quality effort in the department. After posting these results, the Quality Control representative posted charts similar to Figure 3 in the office of each General Foreman showing the results of the previous day's activities for each of his foremen.

Similar charts recapping the operations of each General Foreman were posted in the office of the Superintendent.

This provided for a continuing emphasis by each level of supervision on quality problems. In fact, as a direct result of this system, it became imperative for the foreman to investigate every deviation immediately in order to be able to answer satisfactorily questions which he knew would be forthcoming. In order to aid him in this area, the inspectors were instructed to notify the foreman as soon as any defective material was found rather than waiting for the normal time lag when the information would be posted on the floor chart. The foreman then investigated each such incident and filled out the Foreman's Deviation Report, Figure 4. In this, he identified the difficulty and then attributed the defect to the responsible area and either to operator or non-operator error. If he attributed the defect to causes outside his own department, he was required to obtain the signed concurrence of the person in charge of that department. This report was then signed off by the General Foreman who would also adjudicate any disagreements that arose. An important portion of the form is a section devoted to action taken to prevent recurrence. Follow up on this by the quality organization contributed considerably to a sizable over-all reduction in the defective material produced.

Quality Control reports were issued monthly as shown in Figures 5 and 6. These showed the departmental standings and also the defect figures chargeable to various causes. By this means, quality improvement effort could be concentrated in areas where maximum returns could be obtained.

In the actual installation of the system, each department was handled in turn and meetings were held with every foreman, appropriate union representatives, and the inspectors to make sure everyone was fully aware of the program and its goals. Although there was naturally a certain amount of distrust initially, as the system operated and more and more departments were exposed to it, a spirit of cooperation developed which contributed significantly to the success of the operation.

The experience described above covers a system which is applicable to almost any job-shop in the country. It provides a reasonably simple and inexpensive statistical control chart system for considering as one what previously were considered widely separate and distinct operations. It is an invaluable index for demonstrating whether or not your job-shop has the basic attitude of "making it right the first time" and provides a technique for reaching and continuing to hold that very desirable goal.

INSPECTION DAILY TOTAL

MACHINE NUMBER	PARTS PROCESSED BY					
	FIRST SHIFT		SECOND SHIFT		THIRD SHIFT	
	INSP. POINTS	DEVI-ATIONS	INSP. POINTS	DEVI-ATIONS	INSP. POINTS	DEVI-ATIONS
326			97	2		
417			116	14		
83			17	0		
205			263	1		
641			121	4		

AREA Lathe

J. Jeorge — INSPECTOR — 2 SHIFT — 3/2/46 DATE

FIGURE 1

MACHINE NUMBER	DATE	1ST SHIFT			2ND SHIFT				3RD SHIFT				
		PB	3/1	3/2	3/15	PB	3/1	3/2	3/15	PB	3/1	3/2	3/15
326	NO. DEFECTS	1.4	1	0		2,0	0	2		2.4	1	—	
	NO. INSPECTIONS		49	101			86	97			51	—	
	TOTAL DEFECTS		1	1			0	2			1	1	
	TOTAL INSPECTIONS		49	150			86	183			51	51	
417	NO. DEFECTS	2.3	3	0		5,0	5	(14)		3.6	(8)	1	
	NO. INSPECTIONS		84	89			67	116			101	28	
	TOTAL DEFECTS		3	3			5	(19)			(8)	(9)	
	TOTAL INSPECTIONS		84	173			67	183			101	129	
83	NO. DEFECTS	6.9	1	0		4.2	1	0		.6	0	0	
	NO. INSPECTIONS		9	19			14	17			13	8	
	TOTAL DEFECTS		1	1			1	1			0	0	
	TOTAL INSPECTIONS		9	28			14	31			13	21	
205	NO. DEFECTS	2.1	3	3		2.5	4	1		1.5	0	—	
	NO. INSPECTIONS		147	188			214	263			204	—	
	TOTAL DEFECTS		3	6			4	5			0	0	
	TOTAL INSPECTIONS		147	335			214	477			204	204	
641	NO. DEFECTS	.05	0	(1)		.4	0	(4)		1.0	0	4	
	NO. INSPECTIONS		178	178			88	121			19	162	
	TOTAL DEFECTS		0	1			0	(4)			0	4	
	TOTAL INSPECTONS		178	356			88	209			19	181	

FIGURE 2

FOREMAN'S DEVIATION REPORT

AREA	MACHINE OR OPERATION NUMBER	DATE	SHIFT	NO. OF DEVIATIONS
Lathe	641	3/2/56	2	4

PART NUMBER AND NAME: 304B627 Shaft

OPERATOR ERROR		NON-OPERATOR ERROR	
CAUSE OF DEFECT	X	CAUSE OF DEFECT	X
FAULTY MACHINE OPERATION		PRIOR OPERATION	
MEASUREMENT ERROR		MACHINE	
POOR SET-UP		PLANNING	
MISREAD PRINT		ENGINEERING	
WRONG TOOL	X	HEAT TREAT	
OTHER (SPECIFY)		TOOLING	X
		MATERIAL	
		OTHER (SPECIFY)	

OPERATOR: 114

SHIFT: 2

SPECIFIC CAUSE OF DEFECT: (I.E. GAGE DESIGN, COMPUTATION OF STACK-UP IN PLANNING, ETC.)
Wrong tool provided

COMMENTS:
Operator asked for special tool at tool crib - wrong one issued - he didn't check it, but used it.

ACTION TAKEN TO PREVENT RECURRENCE:
Discussed with operator, tool room supervisor, and tool room attendants.

DISPOSITION OF PART(S):
Scrap.

FOREMAN AND SHIFT AT DEVIATION LOCATION	FOREMAN AND SHIFT RESPONSIBLE	GENERAL FOREMAN'S APPROVAL
J. Severs 2-Lathe	H. Brown - 2 Tool Room	P. M. Sloan

FIGURE 3

QUALITY CONTROL CHART

MONTH March

SHIFT	P8	DATE	3/1	3/2	3/3	3/4	3/5	3/6	3/	/26	3/27	3/28	3/29	3/30	3/31	
1	2.2	NO. DEFECTS	8	4												
		NO. INSPECTIONS	467	575												
		TOTAL DEFECTS	8	12												
		TOTAL INSPECTIONS	467	1042												
2	2.8	NO. DEFECTS	10	21												
		NO. INSPECTIONS	469	614												
		TOTAL DEFECTS	10	31												
		TOTAL INSPECTIONS	469	1083												
3	2.1	NO. DEFECTS	9	5												
		NO. INSPECTIONS	385	198												
		TOTAL DEFECTS	9	14												
		TOTAL INSPECTIONS	385	586												
Total	2.3	NO. DEFECTS	27	30												
		NO. INSPECTIONS	1,324	1,387												
		TOTAL DEFECTS	27	57												
		TOTAL INSPECTIONS	1,324	2,711												

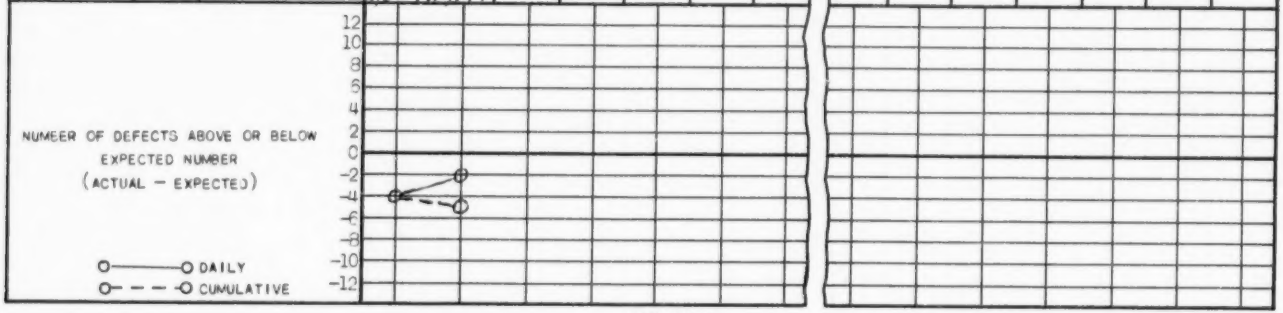


FIGURE 4

QUALITY CONTROL CHART

SHIFT 2 MONTH March

AREA	Fz	DATE	3/1	3/2	3/3	3/4	3/5	3/6	3/	26	3/27	3/28	3/29	3/30	3/31	
LATHE	2.3	NO. DEFECTS	27	30												
		NO. INSPECTIONS	1,324	1,387												
		TOTAL DEFECTS	27	57												
		TOTAL INSPECTIONS	1,324	2711												
MILLS-DRILLS	3.7	NO. DEFECTS	41	4												
		NO. INSPECTIONS	864	162												
		TOTAL DEFECTS	41	45												
		TOTAL INSPECTIONS	864	1026												
SHEET METAL	10.4	NO. DEFECTS	49	44												
		NO. INSPECTIONS	326	465												
		TOTAL DEFECTS	49	93												
		TOTAL INSPECTIONS	326	791												
TOTALS	4.1	NO. DEFECTS	290	271												
		NO. INSPECTIONS	7,219	6,437												
		TOTAL DEFECTS	290	561												
		TOTAL INSPECTIONS	7,219	13,656												

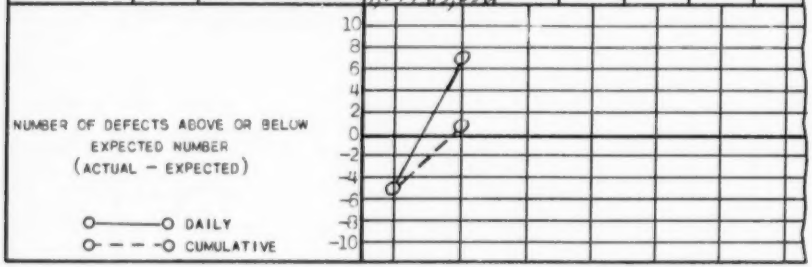


FIGURE 5

QUALITY CONTROL CHART

SHIFT 2

MONTH March

CAUSE	PB		DATE	3/1	3/2	3/3	3/4	3/5	3/6	3/	/26	3/27	3/28	3/29	3/30	3/31	
OPERATOR	6.8	NO. DEFECTS		229	188												
		TOTAL DEFECTS		229	417												
MACHINE	4.6	NO DEFECTS		25	67												
		TOTAL DEFECTS		25	42												
PLANNING	.45	NO. DEFECTS		4	0												
		TOTAL DEFECTS		4	4												
TOOLING	.8	NO. DEFECTS		0	1												
		TOTAL DEFECTS		0	1												
VENDOR	5.9	NO. DEFECTS		31	15												
		TOTAL DEFECTS		31	46												
ENGINEERING	.6	NO. DEFECTS		1	0												
		TOTAL DEFECTS		1	1												
TOTAL	4.1	NO. DEFECTS		290	271												
		NO. INSPECTIONS		7,219	6,437												
		TOTAL DEFECTS		240	561												
		TOTAL INSPECTIONS		7,219	13,656												

NUMBER OF DEFECTS ABOVE OR BELOW EXPECTED NUMBER (ACTUAL − EXPECTED)

10 8 6 4 2 0 -2 -4 -6 -8 -10

O——O DAILY

O- - -O CUMULATIVE

FIGURE 6

QUALITY CONTROL'S OBLIGATION TO MANAGEMENT AND CUSTOMER ON RELIABILITY OF COMPLEX WEAPONS

H. O. Williams

Douglas Aircraft Co., Inc.

The subject of Quality Control's obligation to Management and the customer on reliability of complex weapons is by no means a new subject, as we in Quality Control in the aircraft industry have been endeavoring to fulfill this responsibility for many years, and during all this time, Quality Control has been vitally concerned with safety, dependability, and service life of the final product delivered to our customers.

Recently we find that someone has coined a new phrase for us. Reliability! What is reliability? Is it any different from what we have been striving for these many years? We believe that Management and customer alike are definitely concerned with the following three basic considerations:

1. They want the best product it is possible to produce.

2. They want the best product possible to produce for the least amount of money.

3. They want the product on time.

Now, how do we in Quality Control meet these responsibilities? Since Quality Control does not design and engineer the product involved, nor does Quality Control manufacture the product involved, our responsibility rests with the assurance that the parts, components, and end items are manufactured to the requirements of the designer's dream and the engineering blueprints and specifications. And, so in order to do this, we are charged with the responsibility of organizing and maintaining an effective and economical quality control system.

This system shall have been planned and developed in conjunction with other planning functions, such as production planning and subcontracting planning, based on complexity of design, interchangeability requirements, and manufacturing techniques. The system shall assure that adequate control of quality is maintained throughout the entire process of manufacture, including packaging and shipping and shall provide a means for the ready detection of significant types of recurring discrepancies together with the corrective action taken.

We are further charged with the responsibility of issuing inspection instructions which shall assure that physical, dimensional, and functional properties which would affect safety, or result in substantial reduction of useability, are listed in these instructions.

We are further charged with the responsibility of maintaining adequate records of inspections and tests performed. These records shall provide evidence that the required inspection operations have been performed and shall indicate the number of rejections in each sample, lot, run, etc., and the reasons for such rejections. The results of interchangeability inspections and tests shall be indicated and all of these records shall be available to both top management

and the customer for review. (In fact, we have some customers who we some times believe are more interested in the records than they are of the finished product.)

All of these responsibilities require us to continually maintain an educational program for in these days of increased manufacturing schedules, you are fortunate indeed, if you can hold on to your experienced personnel, and you have to be content with what you can get in the way of new personnel. We must be forever on the alert for engineering and customer changes to insure they are accomplished at the proper time and place. This is very important and can be some times quite costly, as I am sure you will agree with me that you never get the same results, both quality and economically wise when a job is assembled and taken apart and changed, than a job that is done correctly the first time.

Another very important obligation to both Management and the customer on the part of Quality Control is to carry on a program of research and development into new ways and means of doing our job more efficiently and economically. Great strides in this direction have been made during the last few years in the development of facilities and procedures in the field of nondestructive testing and statistical quality controls.

One prominent factor in the success or failure of any organization is the spirit of cooperation that exists between departments and divisions of your company. Many excellent programs and efforts have failed because of the lack of cooperation between people and departments. Quality Control personnel having a close relationship with most of the departments in your company, particularly manufacturing, can do much to encourage and foster a desire to accomplish a better job.

Quality Control should be prepared to convey information to the responsible agencies such as the design engineer or any other agency in your facility responsible for the correlation of the reliability of the end product. This we could be in a position to do by referring to our historical records to provide the design engineer with information pertaining to the in-tolerance variations that the information so gathered may be correlated to end item performance.

THE PROPOSED MILITARY STANDARD FOR INSPECTION BY VARIABLES /1

Harry M. Rosenblatt and William Wolman
Office of Naval Research and Bureau of Yards and Docks
Department of the Navy

Introduction: Origin of Standard

The preparation of this Standard is a natural consequence of the increasing recognition by Industry, University and Government workers in the field of quality control of the importance of inspection by variables for percent defective, as a test procedure to evaluate product quality. In the published and unpublished literature on quality control there is ample evidence of the contributions to statistical research in support of this test procedure and of the numerous applications of the test procedure by various activities. Some activities in and out of government have already prepared sets of standard sampling plans designed to their own preferred methods of application. The difference in methods of application is undoubtedly influenced by the availability of theoretical results at the time of preparation or by different opinions as to ease of application. Nevertheless, these differences are all superficial since, basically, all percent defective inspection by variables sampling plans worthy of the name are cut from the same cloth. In this background, the need to prepare a military standard for inspection by variables for percent defective, to serve as an alternative to Military Standard 105 for inspection by attributes, became apparent.

Accordingly, the Office of the Assistant Secretary of Defense (Supply and Logistics) brought together, as a team, representatives of the various agencies of the Department of Defense to specify the scope and content of the new military standard. This specification included types of sampling plans, lot sizes, sample sizes, acceptable quality levels, risks of incorrect decision for AQL product and the relationship among the various types of sampling plans.

A desire of this task team to utilize the recent theoretical solution, which provided an acceptance region for the two-sided test in inspection by variables comparable to that for the one-sided test, led to a revision of the original acceptance criteria of the solution into a simpler practicable form. Further discussion of this solution will appear later in the paper. The theoretical solution and the revision of the acceptance criteria was accomplished by members of the Applied Mathematics and Statistics Laboratory, Stanford University. This group also completed the sampling procedures and tables within the framework indicated by the Department of Defense Task Team in the form of a first draft and is also responsible for the theoretical basis of the Standard. The Department of Defense preparing activity for the Military Standard is the Bureau of Ordnance, Department of the Navy. This activity prepared several revisions of the Standard and has the task of obtaining final concurrence among the three services in its issue. In addition, the Bureau of Ordnance is completing a report on the full theoretical basis of all the tables and procedures in the Standard. A listing of all those people who worked on or influenced the contents of the

/1 Paper based on work done while both authors were employed by the Bureau of Ordnance, Department of the Navy

Standard in its present form would be long indeed. It is fair to say, however, that they represent leading persons in the fields of quality control and statistics from University, Industry and Government.

Brief Review of Inspection by Variables for Percent Defective as a Test Procedure: If the quality of a product depends on a characteristic, which can be measured on a continuous scale, such as hardness, time, diameter, weight, etc., it is possible to use inspection by variables as a test procedure. If the measure of quality is percent defective there is specified an upper specification limit on the measurable characteristic or a lower specification limit and sometimes both limits. Any unit of product for which the measurement of the quality characteristics exceeds the upper limit and/or is less than the lower limit is a defective unit. The lot percent defective, say p', is then the result one would obtain if he divided the total number of defective units in the lot by the lot size. The objective of the test procedure applied to acceptance sampling is to select a sample from the lot and decide to accept or reject the lot based on the collective measurements of the sample units.

To develop a rational basis for the decision procedure, one specifies some value of lot percent defective, say p'_0 as an acceptable quality level (AQL) and prefers to decide to accept if the true lot percent defective does not exceed this AQL value; likewise some value of the lot percent defective, say p'_1 where $p'_1 > p'_0$, representing an undesirable measure of quality is specified for which it is preferred to decide to reject if the true lot percent defective is equal to or greater than this value. These preferences are made specific by designating suitable risks α and β where α is the risk of making an incorrect decision to reject when the true lot percent defective is p'_0 and β is the risk of making an incorrect decision to accept when the true percent lot defective is p'_1. These risks represent the probability of an incorrect decision, or what is equivalent, the frequency in the long run of making an incorrect decision. An objective of the inspection by variables test procedure is to provide a method for making acceptance decisions which will be controlled within the risks specified for the quality levels indicated. The operating procedures provide the size of the sample to select and an acceptance constant; by comparing the sample results with the acceptance constant, a decision to accept or reject the lot is made. The procedure makes it possible to compute the probability of acceptance for any value of the true lot percent defective, that is, the operating characteristics of the sampling plan. In this Standard, computation of these probabilities is based on the assumption that the quality characteristic is normally distributed.

Coverage of Standard

The Standard is divided into four separate sections as follows: Section A - General Description of Sampling Plans, Section B - Variability Unknown - Standard Deviation Method, Section C - Variability Unknown - Range Method, Section D - Variability known.

Section A, the general description of the sampling plans, will always have to be used in conjunction with the other sections since it provides general concepts and definitions needed for sampling inspection by variables. It includes such topics as Scope of Standard, Acceptable Quality Levels, Submittal of Product, Lot Acceptance, Sample Selection and Operating Characteristic Curves.

Sections B and C describe the specific procedures for application of the sampling plans when variability is unknown. For most applications the variability is unknown and either Section B or C will be used. The term variability is a general one and refers to the measure of dispersion of the quality characteristic for the lot or process. When variability is unknown, it is estimated either by the sample standard deviation under Section B or the average range under Section C. The range method requires, for most of the comparable plans, more observations than the standard deviation method, but has the advantage in being simpler in the computations required.

Section D is available if the variability is known; substantial savings in sample size over unknown variability plans for the same operating characteristics can be made in this case, however, the user is reminded that the requirement of known variability is a stringent one. From a practical viewpoint, variability may be considered "known", if it is observed to remain stable over a period of time, that is, be in "statistical control".

Like MIL STD 105, the sampling plans are indexed by lot size, inspection level and AQL. There are 17 lot size intervals, each identified by a sample size code letter, grouped into 5 inspection levels. Sampling plans are provided for 14 AQL values ranging from .04 to 15.0 percent. Furthermore, the risk involved in rejecting material for a given AQL decreases as the lot size increases, this again follows a pattern similar to that of MIL STD 105.

The corresponding plans in Sections B, C and D were matched as well as possible, under a system of fixed sample sizes, with respect to their operating characteristic curves and are extremely close, in most cases. Hence, for a given lot size, inspection level and AQL, the user may select a plan from either Section B, C or D, and be assured of the same probability of rejecting or accepting material for any given quality; it is important to realize the provision that the variability has to be known for the Section D plans.

Each of the Sections B, C and D consists of three parts:

(1) Acceptance sampling plans for single specification limit case.
(2) Acceptance sampling plans for the double specification limit case.
(3) Procedures for estimation of process average and criteria for tightened and reduced inspection.

The single specification limit case is encountered when the quality requirement is given as either an upper or a lower limit for the characteristic measured and a corresponding single AQL is specified. In this case, the acceptance criterion is given in two forms. For the same sample size, one form provides acceptance constant k similar to that used in the $\bar{X}$-ks procedure; the other form provides a maximum allowable percent defective M and requires estimates of the lot percent defective. Either form may be used, since for the single specification limit case they provide the identical operating characteristics.

Under the double specification limit case, the quality requirement is given by an upper and a lower specification limit for the characteristic measured. If a separate AQL is assigned to each limit with no specific control over the total percent defective outside both limits,

then the procedures for a single limit case can be applied to each limit separately. This is not ordinarily considered to be a double specification limit case.

In most applications of a double specification limit, one AQL is assigned to the combined percent defective above the upper limit and below the lower limit; in certain cases separate AQL values are assigned to the upper and the lower limit and some control is specified over the total percent defective. Procedures for both the single and the separate AQL situation for the double specification limit case are provided in each of the Sections B, C and D. Additional discussion of the problems encountered under this case is given later in the paper.

Estimation of process average under procedures of the Standard is accomplished by simply averaging the estimated lot percent defectives for a number of lots. Procedures for tightened and reduced inspection based on the inspection results of the previous 5, 10 or 15 are also included.

Distinguishing Features: A distinguishing feature of the Standard is the introduction of acceptance criteria by comparing the estimate of the percent defective p as computed from the sample results with a maximum allowable percent defective M obtained from the tables in the Standard. Lot disposition then consists of the following:

If $p \leq M$ accept the lot
$p > M$ reject the lot.

The maximum allowable percent defective M is an acceptance constant which depends on the lot size, inspection level and AQL which is applicable to the situation at hand. In the single specification limit case, p is an estimate of the percent defective above the upper or below the lower specification limit respectively.

This acceptance procedure is analogous to attribute inspection under Military Standard 105 where we compare the number of defectives found in the sample with the acceptance number applicable. We can set up the following comparison in a specific example.

EXAMPLE

Lot Size 3000, Inspection Level II, AQL 1%

	Attribute Inspection under MIL-STD-105	Variables Inspection Under Proposed Standard
Sample size code letter	L	L
Sampling plan	Single	Variability Unknown - Standard Deviation Method
Sample size	n = 150	n = 40
Acceptance constant	$a = (\frac{a}{n} = 2.67\%)$	M = 2.71%
Sample results used	d = number of defectives in sample	p = estimate of lot percent defective based on $\bar{x}$ and s
Acceptance criteria	Compare d and a	Compare p and M
Lot disposition: Accept	If $d \leq 4$ $(\frac{d}{n} \leq 2.67\%)$	If $p \leq 2.71\%$
Reject	If $d > 4$ $(\frac{d}{n} > 2.67\%)$	If $p > 2.71\%$

The percent defective in the lot is estimated by computing a quality index, Q, from the sample results. An example of the quality index for the standard deviation method and a single specification limit, U, is:

$$Q = \frac{U - \bar{X}}{s}$$

where $\bar{X}$ and s are the sample mean and standard deviation, respectively. The value Q is then used to read off the estimated percent defective, p, from a table provided in the Standard. The estimate p obtained in this manner is the best possible in the sense that it is the unique unbiased minimum variance estimate of the lot percent defective.

Most important, however, in sampling inspection by variables is the fact that the use of this estimate p provides a solution to the double specification limit problem when a single AQL is specified. This problem consists of providing an acceptance procedure for the double specification limit case so that the operating characteristics are the same as that for the single specification limit case. To illustrate what is involved, Chart I shows a normal distribution curve with the upper limit indicated by U and the lower limit by L. The lot percent defective above U and below L are indicated by p'_u and p'_l respectively. The mathematical procedures necessary to compute the exact OC curve require knowledge of p'_u and p'_l separately and not their sum. Hence, there is a separate OC curve for each split-up of total percent defective p' which equals $p'_u + p'_l$. The procedures of the Standard using the unbiased minimum variance estimate p are such that, for practical purposes, almost identical OC curves are obtained for a fixed value of

p' regardless of this split. No other known procedure has this property with such exactitude. The Operating Characteristic (OC) curve of a sampling is shown as Chart II. Similar curves are given for all the plans in the Standard. It should be noted that the curves are based on the assumption that the measured characteristic of the individual items follow a normal distribution. The curves represent the percent of lots expected to be accepted for given percent defective.

We had suggested an alternate procedure before the theoretical basis for the present two-sided test was developed. That procedure provided a maximum allowable standard deviation in conjunction with two one-sided tests as the acceptance criteria and resulted in an adequate approximation to the one-sided test for many applications. However, we now consider that it has served its purpose and is now unnecessary as an acceptance criteria in view of the superiority of the now available two-sided procedure which is incorporated in the Standard. Nevertheless, the idea of a maximum standard deviation may have some value to users of the Standard in indicating an upper bound for sample values of the standard deviation. For this reason, a procedure for computing the maximum standard deviation (MSD) has been incorporated in the Standard. One can feel that sample values of the standard deviation less than the MSD help insure but do not guarantee lot acceptance.

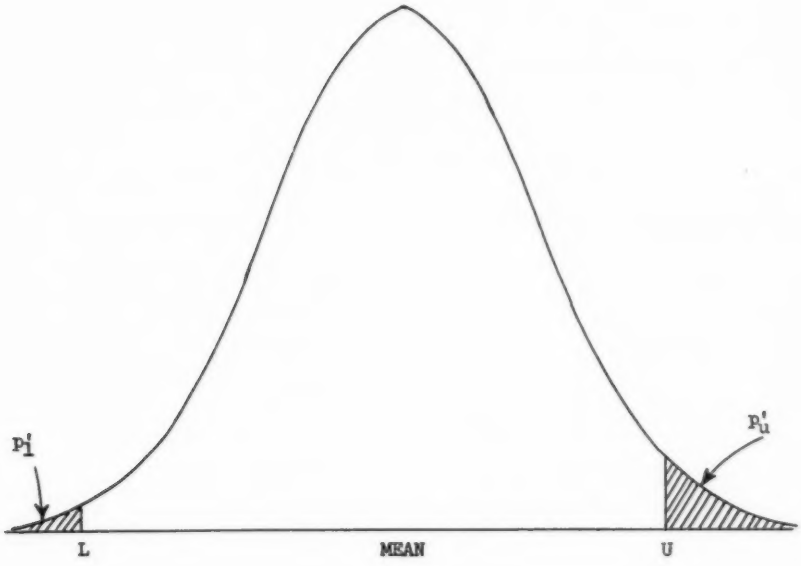
For the future one can envision extension of the procedures of the Standard for combining the results of several characteristics inspected by variables and also combining percent defectives obtained from attribute and variables inspection. Research for this purpose is presently underway.

Chart III gives a detailed example of the procedure for a double specification limit case using the standard deviation method.

Chart I

EXAMPLE OF DOUBLE SPECIFICATION LIMIT

Normal Distribution of Items



U = Upper Specification Limit
L = Lower Specification Limit
p_u' = Percent of items above U
p_l' = Percent of items below L
p' = $p_u' + p_l'$ = Percent of items outside of limits

CHART II

TYPICAL OPERATING CHARACTERISTIC CURVE

AQL = 1%, Lot Size = 3000, Sample Size Code Letter L

_____Variables Sampling Plan - Standard Deviation Method:
Sample Size (n) = 40
Maximum Allowable
Percent Defective (M) = 2.71%

- - - - Attribute Sampling Plan - Single Sampling:
Sample Size (n) = 150
Maximum Allowable
Number of Defectives (a) = 4

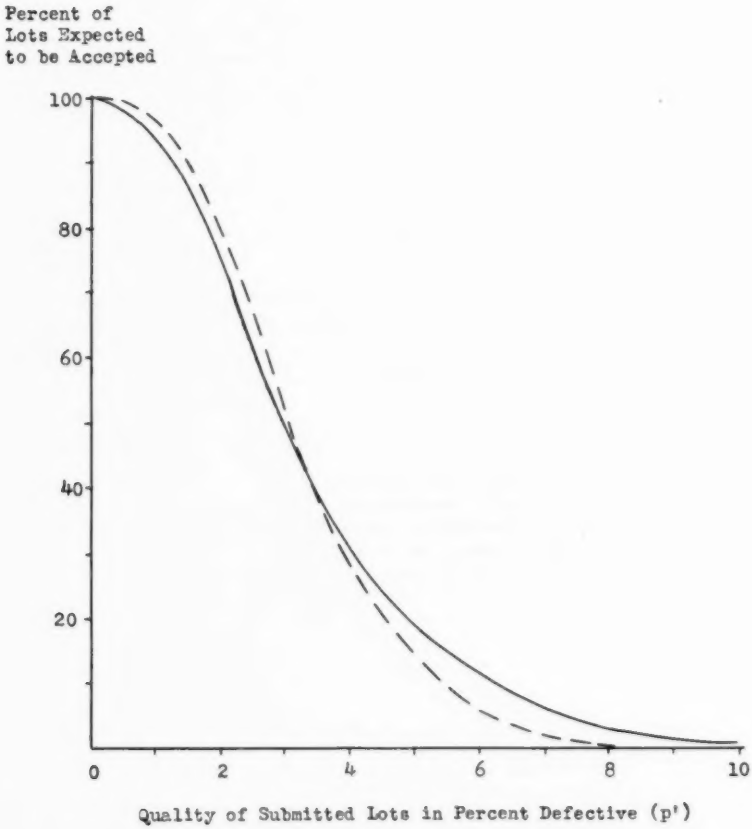


CHART III

Example of Calculations Double Specification Limit

Variability Unknown - Standard Deviation Method
One AQL Value for both Upper and Lower Specification Limit Combined

Example: The minimum temperature of operation for a certain device is specified as 180°F. The maximum temperature is 209°F. A lot of 40 items is submitted for inspection. Inspection level II, normal inspection, with AQL = 1% is to be used. From the Standard, it is seen that a sample of size 5 is required. Suppose the measurements obtained are as follows:

197°, 188°, 184°, 205° and 201°
and disposition of the lot is to be determined.

Line	Information needed	Values Obtained	Explanatio
1	Sample Size: n	5	
2	Sum of Measurements: ΣX	975	
3	Sum of Squared Measurements: ΣX^2	190,435	
4	Correction Factor (CF): $(\Sigma X)^2/n$	190,125	$(975)^2/5$
5	Corrected Sum of Squares (SS): ΣX^2CF	310	190,435-190,125
6	Variance (V): SS/(n-1)	77.5	310/4
7	Estimate of Lot Standard Deviation(s): $\sqrt{V}$	8.80	$\sqrt{77.5}$
8	Sample Mean ($\bar{X}$): $\Sigma X/n$	195	975/5
9	Upper Specification Limit: U	209	
10	Lower Specification Limit: L	180	
11	Quality Index: $Q_U = (U-\bar{X})/s$	1.59	(209-195)/8.80
12	Quality Index: $Q_L = (\bar{X}-L)/s$	1.70	(195-180)/8.80
13	Est. of Lot Percent Def. above U (p_u)	2.19%	From Standard
14	Est. of Lot Percent Def. below L (p_l)	.66%	From Standard
15	Total Est. Percent Def. in Lot (p): $p = p_u + p_l$	2.85%	2.19%+.66%
16	Max. Allowable Percent Def. (M)	3.32%	From Standard
17	Acceptance Criterion: Compare $p = p_u + p_l$ with M	2.85%(3.32%)	From Standard

Lot Disposition. Accept or Reject: Accept the lot, since $p = p_u + p_l$ is less than M.

THE SCIENCE OF "QUALITY CONTROL AS AN OPERATION"

Paul S. Olmstead
Bell Telephone Laboratories
Whippany, New Jersey

1. Introduction

For the last three years, each newly elected Fellow of the American Society for Quality Control has received a letter of announcement from the Chairman of the Examining Committee to the following effect:

"Your election is in recognition of your important contribution to the Science of Quality Control."

The letter then proceeds to mention the specific accomplishments on which the election was based.

The first Constitution of the American Society for Quality Control (page 5, March 1946, Industrial Quality Control) gives the following as the first purpose of the society:

"The purpose of this Society shall be the advancement of and diffusion of knowledge of the science of Quality Control and its applications to industrial progress."

In the latest Yearbook, 1955-6, page iii, the Constitution as amended to June 9, 1954 states:

"Section 1.b. The purpose of this Society shall be:
(1) To create, promote and stimulate interest in the advancement and diffusion of knowledge of the science of quality control and of its application to industrial processes."

It is quite obvious that the inclusion of the phrase, "Science of Quality Control," has not been fortuitous.

According to Webster's Unabridged Dictionary, science is defined in several ways:

"Any branch or department of systematized knowledge considered as a distinct field of investigation or study," or

"A branch of study which is concerned with observation and classification of facts, especially with the establishment of verifiable general laws, chiefly by induction and hypothesis," or

"Specifically, accumulated and accepted knowledge which has been systematized and formulated with reference to the discovery of general truths or the operation of general laws; knowledge classified and made available in work, life, or the search for truth."

It is quite apparent that Quality Control qualifies as a science by these definitions. It is a distinct field. It is concerned with establishing verifiable general laws. Also, its future is based on present

knowledge systematized and formulated to facilitate the discovery of new truths or of the operation of new laws.

One difficulty in discussing Quality Control is the fact that its meaning to a particular individual depends on his experience in the Quality Control field. In general, however, the members of the American Society for Quality Control are interested in the broadest meaning of the term. They accept Quality Control as an operation for assuring the best quality for the least cost. This includes as a major part the "operation of statistical control," first described by Walter A. Shewhart in his book, "Statistical Method from the Viewpoint of Quality Control," Graduate School, Department of Agriculture, Washington, D. C., 1939. A summary of recent developments in the identification of Quality Control as an Operation has been provided by Miss Mary N. Torrey in a review article, "Quality Control in Electronics," published in the IRE Proceedings, November 1956.

In accepting Quality Control as an Operation for assuring the best quality for the least cost, it is assumed that it is being applied to a system determined by the laws of several other sciences such as Physics, Chemistry, Psychology, Economics, etc. It may be of interest, therefore, to trace the change in assumptions associated with various sciences in order to reach the goal of the "best quality for the least cost." First, an ideal science will be assumed.

2. Quality in an Ideal Science

In an ideal science, all constants of nature are in fact constants. All independent variables are known and may be determined exactly. Under these conditions, it is customary to consider the quality of a thing as a characteristic or set of characteristics of the thing. Mathematically, this may be expressed as follows:

$$Q = f(x_1, x_2, x_3, \ldots\ldots x_n) \qquad (1)$$

Thus, the Quality, Q, is considered to be a function of the n independent characteristics or variables, x. It is therefore true that each set of x's determines the value of Q exactly. The various theorems of mathematics may be used to illustrate this definition of quality.

3. Quality in an Applied Science

In a theoretical or ideal science, no difficulty is encountered with equation (1). However, in applying a science, experiments are made to check the equation. Since the agreement is not always exact, a modification must be made:

$$Q_i = f(x_1, x_2, \ldots\ldots x_n) + e_i \qquad (2)$$

Here, the function, f, is the same as in equation (1), Q_i is the value of Q observed in the i^{th} experiment, and e_i is the difference, $Q_i - Q$, observed in that experiment. In some experiments of this type, the experimenter found that, if he made a large number, the average of his

e's became quite small, sometimes close to zero. As a result, he attributed his e's to errors of measurement. In other cases, he found that his average was based on all positive (or all negative) e's. This was attributed to systematic error or bias. At other times, the range of observed e's changed with each series of experiments.

From such observations, it is clear that, to reach a point of complete determination of Q in each and every experiment, provision must be made for unknown x's in addition to the n x's of equation (2). If it is assumed that the total x's for such a determination is finite, say N, it is possible to write the following theoretical expression for Q:

$$Q = f(x_1, x_2, \ldots\ldots x_n, x_{n+1}, \ldots\ldots\ldots x_N) \qquad (3)$$

Thus, if all N x's were known, Q could be determined without error. However, when only n x's are known, the practical equation is (2) above.

In applied science, some qualities are established on an empirical basis before a theoretical basis has been discovered. In these instances, it is desirable to determine an equation relating Q to the known x's. Equation (2) also meets the requirements for this case but with this difference. The function, f, is not known but must be determined from the data. In the simplest case, this means determination of the coefficients in a linear equation in a way that will minimize the contribution of the e_i's. The method of least squares has often been used for this purpose. However, in more complex experiments where replications are made, analysis of variance is now used to provide estimates of the "best" constants to be derived for the function, f.

Recently, the notion of team research is being advocated more and more for applied science. This aims to bring together people from several science fields to work on a single complex project. Obviously, this broadens the base for determining quality. This will now be considered.

4. Quality in Operations Research

The team approach was given the name of Operations Research during World War II. Essentially, it resulted in bringing together experts in various fields and thereby increased the number of known x's from the n in one applied field to n+m, the number based on the combined knowledge of the team. Under these conditions, it is now possible to write the equation for Q_i as:

$$Q_i = f(x_1, x_2, \ldots\ldots x_n, x_{n+1}, \ldots\ldots x_{n+m}) + e_i \qquad (4)$$

This has the same form as equation (2) but provides for the additional m x's that result from combining the knowledge relating to several fields.

Operations Research has done a great deal about bringing workers in various sciences together. In particular, it has popularized many specialized techniques such as queuing theory, linear programming and

the like. Specifically, it has placed more emphasis on cost and value as items contributing to quality. However, the situation with respect to equation (4) is the same as that for equation (2). It is still necessary to determine whether an assumed function, f, is satisfactory to describe the observed values of Q, or conversely to estimate the "best" values for the constants in the function, f, from the observed values of Q. The commonly accepted procedure is that mentioned in the previous section.

5. Scientific Method

A term frequently used in science is scientific method. Ordinarily, this is considered to be a three step process:

Hypothesis

Experiment

Test of Hypothesis

Hypothesis is a statement of what may be true. Experiment is obtaining data for the purpose of proving that the hypothesis is true or false. Test of Hypothesis is using the data to show whether the hypothesis is true or false.

There have been two important motivating factors that have led to developing scientific method:

Improvement in Precision of Measurement

Desire for Extrapolation

In each case, the question must be raised: Will accepted theory hold? In many instances, it has not and new theory has had to be derived or invented. This is essentially the scientific method of discovery, a search for nonconformance.

6. Quality in the Science of "Quality Control as an Operation"

One of the primary assumptions of Quality Control is that the function, f, is never sufficiently well known at the start of any investigation or repetitive experiment to serve as a complete description of the expected data. More specifically, it is almost certain that the observed data will fail to meet requirements for statistical control. Further, Quality Control procedure will ordinarily lead to the discovery of Assignable Causes of variation, i. e., new x's. This means that the Quality Control procedure will detect the presence of an unknown x, relate it to the conditions under which it occurred, determine the magnitude of its effect, and thereby assist in its discovery and identification. This suggests the following modification in the previous definition of Q:

$$Q = f(x_1, x_2, \ldots x_{n+m}, x_{n+m+1}, \ldots x_{n+m+k}, \ldots x_N) \qquad (5)$$

This is equation (3) modified to include the knowledge inherent to Operations Research and to indicate that some of the remaining x's, say k, may be identified through the operation of Quality Control.

The first step in a Quality Control analysis is to determine equations (4). Ordinarily, these would be for the case where the "best" estimates of the constants in the function, f, have been inserted in the equations. The next step in a Quality Control analysis is to examine the e's for statistical control. Stated as an hypothesis, this may be written in the form:

$$e_i = f(x_{n+m+1}) + \Delta e_i \qquad (6)$$

This postulates the existence of an undiscovered Assignable Cause, x_{n+m+1}, that may be taken into account to reduce the e's to Δ e's. The simplest form of this type of analysis is the control chart. It is well known that it is relatively easy to identify Assignable Causes in this way.

The Box-Wilson Method of Response Surfaces, Jour. Roy. Stat. Soc., Ser. B, 13, pp 1-45, 1951, is a more complicated procedure that also has as its aim the solution of equation (6). It introduces regression analysis and specific experimental designs useful in determining the coefficients of new x's that are based on second order effects of the x's. This method does not provide for new x's that are not related to the old. However, it does provide an important first step in finding new x's.

Discovery of new x's by applying Quality Control criteria to equation (6) permits redetermination of equation (4) to include the new x's. Up to this point, each x has been assumed to have a fixed value. In any important practical case, the value of each x will change with each experiment, i, even when the x is in theory being held constant. This leads to the following for Q_i:

$$Q_i = f(\bar{x}_1, \bar{x}_2, \ldots\ldots \bar{x}_{n+m+k}) + E_i \qquad (7)$$

where the E's include the effects of minor changes in the x's as well as the e's, and the k newly discovered x's are included in the equation. The $\bar{x}$'s in this equation are used to indicate that the x's are held "constant" in the experimental sense.

7. Control in the Science of "Quality Control as an Operation"

Equation (7) is the basic equation for control in the science of Quality Control. Essentially, it states that, if each of the x's is controlled, Q will be controlled. Thus, the expected value of Q and its expected variability can be predicted from the known average x's and their variabilities. In a production process, this means that the design quality, $\bar{Q}$, of the product is related to the design qualities, $\bar{x}$'s, of the subassemblies, piece parts, and raw materials. However, the variability in Q is dependent on the variabilities in the x's as they are being produced for assembly. It is important, therefore, that

deviations from control be discovered at the raw material, piece part, or subassembly stage rather than in the final assembly or later in order to assure satisfactory quality.

Such discovery and correction is part of the Operation of Quality Control. Discovery in the simplest case is based on the use of control charts to identify points out of control. Correction is obtained by feeding back information relating to these points to the individual responsible for the process. His action then will be based both on his knowledge of the process and the new information given him. Ideally, this action is often taken as a team responsibility with representatives of design, manufacture, and quality control serving on the team.

In research and development work and also in field trials of new products, an understanding of equation (7) is also important. Essentially, it states that, if the basic factors, the x's, can be isolated, a good estimate of Q and its expected variability may be made. In one instance, a prediction determined from the performance of the first 800 units of a new product was checked ten years later with a trial of 1200 units. In the meantime, some 40,000,000 units had been made with the usual minor modification from time to time. The new trial showed that the original prediction was still good. As additional verification that the original prediction had been sound, separate comparison was made of prediction and performance for each of the seven major components. All were good.

It seems quite likely that the principle of control should be useful in studying various economic phenomena such as those related to growth. It is reasonable that equation (7) would be useful here. The problem is simply to identify the primary causal factors, x's, establish the "best" estimates of the constants for the coefficients, and then study the E's for control.

Standardization, that is, development of standards, is basic to all science and is becoming more and more important in international trade. It is worth noting that the operational meaning of a standard depends on the concept of control. This was first pointed out by Shewhart in his book, "Statistical Method from the Viewpoint of Quality Control."

8. Unsolved Problems in the Science of "Quality Control as an Operation"

An important characteristic of a science is its unsolved problems. Some relate to theory and others to application. Some await the carrying out of new experiments. Others may be solved by reconsidering data already available. A few that will require continuing attention in the Science of "Quality Control as an Operation" are the following:

a. Development of new statistical techniques for detecting particular types of assignable causes

b. Determination of the practical efficiency associated with the use of particular statistical techniques for detecting particular types of assignable causes

c. Modification of the techniques in such fields as Operations Research, Cost Accounting, and many others to include the concept of control

d. Organization of Quality Control itself in the broad sense of introducing the concept of control in determining what is wanted, performing research and development work, designing and specifying the product, making it, inspecting it, and testing it in service to see that it satisfies the wants of the user

These are of the nature of continuing problems for which new solutions will be found from time to time. From an operational view, each new solution represents a discovery that provides a new starting point for reconsidering the problem. It is the application of the concept of control to the science of quality control itself.

9. Summary

This paper outlines the development of some basic concepts in the Science of "Quality Control as an Operation." Particular attention is given to the meaning of Quality which is shown to have an operationally verifiable value only when it is based on data that meet a criterion of control.

The Operation of Quality Control is indicated as a self corrective one aimed at discovering new factors, Assignable Causes, that will permit defining Quality within limits to meet economic objectives.

Mention is made of some of the continuing unsolved problems in this new science. It is indicated that their solution will itself involve the use of "Quality Control as an Operation."

THE QUALITY MANAGER AND QUALITY COSTS

W.J. Masser
General Electric Company

The operation of every business is based on costs. There are costs for Marketing, Engineering, Manufacturing, along with other costs that are necessary to design, to make, to sell and ship to the Market Place, products which influence customer desires.

Interwoven throughout all these costs are the important expense that are devoted to acquire quality components, to produce a satisfactory product and to constantly maintain the desired product quality levels.

My remarks today, will be devoted mainly to the costs related to assure product quality - because all industry is presently plagued with continuing upward pressure by their customers for better and better quality products. And, this pressure comes at a time when quality costs are at their highest level, while at the same time many businesses are operating with obsolete quality practices.

The challenge to gain customer satisfaction, while at the same time reducing the present high cost for quality, rests with the Managers of Quality Control.

To meet this challenge, the Quality Control Manager must become adept in analyzing the whole business cycle of his business from the standpoint of quality effort and quality costs. To do this there is a positive need for establishing quality objectives, quality goals and needed quality programs that will assure the success of any planned Quality Control activity. Along with such objectives, goals and programs is the need for measuring the progress of the Quality Control effort and the extent to which Quality Control makes a contribution to the business as a whole.

As a first approach, to aid Quality Control Management, certain questions must be answered - such as,

1. What are our present Quality Levels?
2. How are we controlling these levels?
3. What does it take in Quality Costs to support this control?

The answers to questions such as these can be, at times, very difficult to obtain without the proper methods. The need therefore is definitely for better measurement techniques.

This need is illustrated by the story of one very ambitious boy who stopped at the corner drug store to use the phone. This is the story that Mr. Smith, the druggist, heard.

"Hello Dr. Jones? Do you need a boy to cut your grass? You don't?

You just hired a boy? Well, how is he doing? Is he giving satisfaction?

He is! Well, thank you, Dr. Jones."

As the boy was leaving the drug store Mr. Smith expressed his sympathy to the boy for not getting the job, to which the boy replied, "Oh, I got the job alright, I was just checking."

And so it is with the Quality Control Manager. He must also check and recheck for good performance. A method that we have found very useful in this area is the Quality Cost analysis method. To briefly explain, Quality Costs, as we see it, are segregated into three distinct areas:

1. Prevention cost is the money expended for the purpose of keeping defects from occurring in the first place. Included here are such costs as quality control engineering, the writing of quality assurance instructions, employee quality training and the quality maintenance of patterns and tools.

2. Appraisal costs which include the expenses for maintaining company quality levels by means of formal evaluations of product quality. This involves such cost elements as inspection, test, quality audits, laboratory acceptance examinations and outside endorsements.

3. Failure costs which are caused by defective materials and products that do not meet company quality specifications. These include such loss elements as scrap, spoilage, rework, field complaints, etc.

There is a definite advantage for grouping the Quality Cost elements into these three major areas - since it simplifies the grouping of twenty four cost elements which make up a complete quality cost analysis record. Therefore, instead of twenty four variables, only three basic variables are involved.

These three major areas now tell us where and how much money we are spending for Prevention, Appraisal and Failure. However, to complete measurements it is necessary to have comparison bases to enable us to relate these quality costs to the rest of the business operation. Here are three comparison bases that have been used successfully in some businesses.

1. Contributed Value.

   This is the value contributed by the product department in designing, manufacturing and selling the end product. It is calculated by subtracting the cost of outside purchased materials and services from Net Sales Billed.

2. Net Sales Billed.

   This is the total amount that is billed, by the department, for products sold during a given period.

3. Operation Labor.

   This is the actual input of dollars for labor and includes good labor as well as labor losses.

Having three bases such as these will assure you that at least one base will reflect a true indication of comparison when the other bases are temporarily thrown out of balance by possible short term changes.

To set up such a quality cost work sheet it is necessary to be sure that the cost accounting methods in a particular business have identified and grouped Quality Costs in a form suitable for the development of adequate control.

Once a system is developed regarding the method to be used in determining a business's Quality Costs and the bases used for comparison purposes, subsequent changes in the original bases obviously invalidates all past data. Therefore, the importance of excercising judgement and caution in selecting these bases cannot be over-emphasized.

To be assured, that the quality costs used for analysis purposes are dependable, the Quality Control Manager must look for three specific things.

The first is that the quality costs are accurate according to the definitions of the elements under the three areas, Prevention, Appraisal, and Failure.

Second, to be sure that the measuring bases; i.e. Contributed Value, Net Sales Billed or Operation Labor, fit the ups and downs of the particular business.

And third, be sure that the quality costs are issued periodically; preferrably on a monthly or quarterly basis.

Now let us look at an example of how the three areas, Prevention, Appraisal and Failure, and their comparison bases, Contributed Value, Net Sales Billed and Operation Labor are related. This will illustrate how this information would appear on a quality cost report. Normally on a standard quality cost report the figures shown would be in dollars. For ease of comparison of the examples shown let us use only percentages instead of dollars and use only two quarters of a whole year.

Example	"X" Department Quality Costs	
	1st Qtr.	2nd Qtr.
Prevention	7.7%	6.4%
Appraisal	23.8%	23.0%
Failure	68.5%	70.6%
Total	100. %	100. %

Total Quality Costs to Measurement Bases:

Quality Cost % to:

Operation Labor	105.0%	106.0%
Contributed Value	10.2%	10.7%
Net Sales Billed	7.8%	8.2%

This example, as I mentioned before, has been illustrated by

showing only two quarters. It is felt that by establishing definite periods of coverage or compiling Quality Costs on a quarterly basis for at least a minimum of six quarters, the most effective analysis can be made. As can be seen from the example of two quarters, the percentages begin to tell a story of the cost of quality. They show the percentage spent in each area, the relation to the comparison bases and the difference between two quarters. However, it is quite apparent, the two quarters also show that there is still not sufficient data to make a decision. Six quarters will give the needed data and will show trends.

To gain additional data, these quality costs broken down by product lines or segments of process flow will allow pin-pointing of the areas that merit the highest priority of quality control effort. To illustrate how this is done and show the results obtained let us break down the 1st quarter into two product lines and relate them to the comparison bases.

By breaking our costs down by product lines makes it possible to see the comparison of one line to another and also highlights additional facts necessary to making a sound decision.

Example V. "X" Department Quality Cost By Product Lines

Areas		1st Qtr. Line A	1st Qtr. Line B
Prevention		5.0%	10.4%
Appraisal		16.2%	32.0%
Failure		78.8%	57.6%
	Total	100. %	100. %
Quality Costs % to:			
Operation Labor		124.0%	86.0%
Contributed Value		11.2%	9.2%
Net Sales Billed		9.3%	6.3%

All of the quality cost data and information is vitally necessary for the Management of Quality Control. There are four major results obtained from accurate Quality Costs.

1. Quality Costs provide an accurate tool for measurement of overall business quality performance.

2. Quality costs are an analysis tool which are used to indicate where quality money is being spent.

3. They provide a tool for programming the when, the where and the how of quality improvements.

4. They make a fine budget tool to enable the forecasting of realistic needs.

These four advantages assist the Quality Manager better to assume his responsibility of maintaining Quality at an optimum quality cost. Since he now can establish realistic objectives, goals and programs to

carry out his job and to obtain the greatest return from his organization through better direction of effort. And even another advantage is the opportunity to set correct product quality levels which will return optimum quality benefits to the business and to the customer.

By utilizing Quality Costs, to their fullest extent, the Manager is able to make his contribution to the businesses, profitability, assist in the realization of greater cost reductions and also help the business maintain a competitive position in the consumer market.

AN OPERATIONAL VIEW OF EQUIPMENT RELIABILITY

R. Drenick
Bell Telephone Laboratories*

Introduction

Some of the recent discussions on the subject of reliability have been pervaded by the uneasy feeling that, while a lot could be said about the subject, no one was altogether certain of just what the subject was. In fact, there was some rather strong indication that not all concerned were talking about the same subject.

The term reliability, it is true, is quite universally used to describe the extent to which a type of equipment is likely to retain desirable performance characteristics. The inconsistency of usage arises when some more quantitative definitions are proposed. There is certainly no harm in having as many definitions of reliability as there are discussants of the subject, provided only that each one of them understands what the other means by his "reliability", and perhaps also why he considers it most appropriate. They can then compare notes in a logical fashion on what their problems have in common and what conclusions and solutions derived by one can be used also by the other.

It is this general topic with which this article is concerned. It develops first a common ground for such discussions by suggesting a fairly general concept of reliability and by exhibiting, through a few examples, that under certain circumstances one reliability definition is in fact more appropriate than most others. Later, an attempt is made to show that care should be taken in drawing conclusions concerning reliability improvements: Procedures, such as preventive maintenance or duplication of equipment, may be highly economical in some operations and yet quite unattractive in others.

The discussion in this paper will utilize examples throughout to make its points. This has the disadvantage of making the argument seem spotty and limited. Unfortunately, it takes considerable attention to detail to treat the subject reasonably completely, more detail in any case than seemed appropriate here. The study on which the examples are based is to be submitted for publication elsewhere.

1. What is equipment reliability?

The impression that the term "reliability of equipment" means different things to different people is readily confirmed by a review of the literature on the subject.

The general practice is, of course, to have reliability be represented by a number which is the larger, the lower the incidence of failure. High favorites for this number are the survival probability of the device (over some given period of time), or its mean life. The opposite is also encountered. That is to say, the number used for this

---

*This paper is based on work done by the writer when employed by the Radio Corporation of America.

purpose decreases with the incidence of failure (such as the maintenance ratio or the replacement rate), in which case the number is understood to represent unreliability. Whenever necessary, reliability is converted into unreliability (or conversely), sometimes by subtracting it from unity, sometimes by taking the reciprocal, whatever seems most appropriate.

There is a way of putting these practices on a somewhat firmer common footing. "Reliability", broadly speaking, is a term used to describe the quality of operation of a type of equipment in regard to deterioration and failure. What is needed, therefore, is a quantitative measure for this, taking due account of the fact that deterioration and failure are random phenomena. The situation is strongly reminiful of the type with which statisticians have been dealing for some time and for which they have found a very satisfying and flexible method of attack: When they need to measure the desirability or undesirability of the outcome of some uncertain event or series of events, they assign a numerical gain or loss to it, and use the mean value of this gain or loss as the required measure.

The same idea can be applied here. To each failure, or series of failures, in a given type of equipment there is assigned a certain gain (or loss) and reliability is defined as the mean gain (or loss) which the operator of this equipment derives from it.

This may seem a very artificial and clumsy concept, and quite possibly it is. In order to assess its usefulness, it should be tested from three points of view:

(1) Is there an easy and natural way of fixing gains and losses in practice?

(2) Under circumstances under which one expects the conventional definitions of reliability to apply, does the one proposed here reduce to these?

(3) Does the present concept shed any new light, or allow any new conclusions, which could not have been easily derived without it?

The discussion which follows is intended to suggest an answer to these questions. After some preliminaries in the next section, a series of typical situations are reviewed which show how gains and losses can be chosen, and that in fact the reliability concepts are obtained which one would intuitively associate with those situations.

2. Preliminaries

Before the main topic of this paper is taken up, a few preliminaries may be useful to establish rapport on the terminology and notation to be used below. The general pattern of failure of a component will be called its "failure law", following what seems rather common practice. Mathematically, a failure law can be characterized in a number of equivalent ways. Most frequently encountered is a characterization in terms of the probability of survival for a period of time t, which shall be denoted here with $F_s(t)$. The failure rate $g(t)$, or the hazard, of a

component is also used, though perhaps not as frequently. It is related to the survival probability by the formula

$$g(t) = -\frac{d}{dt} \log F_s(t).$$

The failure laws which are much used in practice are the exponential and the normal. The first of these is characterized by

$$F_s(t) = \exp\left(-\frac{t}{m}\right), \quad g(t) = \frac{1}{m}$$

where m is the mean life of the component. The failure rate is seen to be a constant, namely, the reciprocal of the mean life. An interesting fact is that the age of a component obeying the exponential law has no bearing on the likelihood of failure.

The normal law has the well-known shape which need not be reproduced here. Less well known may be the corresponding failure rate which is shown in Fig. 1, for several values of m/σ. The normal failure rate invariably rises to infinity as the age t of the equipment increases, a fact which it shares with probably the failure rates of all physical equipment.

3. Special cases

In this section, a review will be made of three examples which, it is hoped, will illustrate the idea of defining "reliability" as a mean operational gain or loss. The first of these is a situation in which a conventional reliability concept, namely, the survival probability, has been patently appropriate. The point here will be to show, for one, that the definition of reliability suggested here confirms the survival probability as the most appropriate, and that the choice of gain or loss to be associated with the operation is a rather natural and straightforward one. The second example leads to a measure of reliability, namely, the replacement rate which is not overly common but which has some interesting properties. The last example deals with a simple situation which has not been as generally studied but which may suggest the usefulness of the reliability concept developed here.

Each example will be given a name, hopefully a reasonably descriptive one, which is intended to characterize but not to limit it. The mathematical detail and its derivation will be omitted in all cases and only results will be mentioned.

Example 1: Guided missile

This example is chosen as the first because it is conceptually the simplest. A guided missile, and for that matter every component of one, is a piece of equipment which is considered a total loss after the first failure in flight. No repairs are possible then.

It has been customary to think of the survival probability $F_s(T)$ for the time of flight T of the missile as the appropriate measure of reliability.

To apply the present concept, a unit gain is assigned to each successful flight, and a zero gain to each unsuccessful one. The mean gain R which is in this case derived, that is the reliability as defined here, is found to be

$$R = F_s(T),$$

that is, again the survival probability.

Example 2: Replaceable component

This example will differ from the previous one in that the device whose reliability is to be investigated is assumed indefinitely replaceable (like, say, a radio tube but unlike the preceding example in which no repair or replacement is contemplated). It will further be assumed that the device is replaced or completely repaired immediately after each failure.

Intuitively, such a component will be considered the less reliable the more often (per unit time) replacements are expected to be necessary. In this example, unlike in its predecessor, therefore, the unreliability is to be characterized, as opposed to the reliability. This can be achieved by assigning a unit loss or, which is the same thing, a gain of minus unity to each replacement and by then calculating the mean loss (or gain) per unit time.

What is obtained by this calculation is a quantity which has been used in practice and which has come to be called the replacement rate. Contrary perhaps to intuition, it is not a very simple quantity: Its derivation from, say, the survival probability of the part can be a fairly tedious affair, and requires the solution of an integral equation, called very appropriately the "renewal equation". This equation has apparently received relatively little attention in recent work on reliability but it has been much studied by actuaries (1) and some of their results may be of interest here.

The replacement rate usually starts out with an oscillation which can be quite pronounced, like some of the ones derived from the normal law, shown in Fig. 2. Invariably, however, the oscillations die down after several mean lives. The resulting plateau is called the "steady-state" of the replacement rate and its level is the reciprocal of the mean life ($1/m$). There is only one failure law for which the replacement rate does not oscillate at all but is constant throughout, and that, as may be expected, is the exponential failure law.

Example 3: Automobile battery

The manufacturers of automobile batteries are in the habit of "guaranteeing" a certain life T for their product. By that they mean that if the battery fails at some earlier time t they will prorate the original purchase price and make a corresponding allowance for the unexpired life towards the purchase of a new battery made by the same manufacturer. That is to say, they will charge for the new battery instead of the list price C a prorated price

$$C - \frac{T-t}{T} C = \frac{t}{T} C.$$

This is, therefore, the gain which a manufacturer derives from a battery replacing another one which has failed at the time t, prior to the guaranteed life.

The reliability of battery, from the manufacturer's point of view, is his average income, that is, the gain per car and per month, say, from his battery sales. If he calculates this he finds

$$R = \frac{m}{T} C - C \left[F_s(T) + \int_T^{\infty} \frac{t}{T} \, dF_s(t) \right].$$

This equation may not be very instructive as it stands but it can be analyzed fairly easily. It develops that the maximum income (per car and per month) which the manufacturer can possibly hope for, and hence the maximum rate of expenditure which the buying public must fear, is C/T; however, he can achieve this only by the miracular feat of building batteries which last for exactly the warranty period and fail immediately thereafter. It is unlikely that he can achieve this but he can seek to approximate it and, according to rumors, this is exactly what battery manufacturers strive to do.

4. Preventive maintenance

The examples which have been discussed above are descriptive: They attach a numerical value to what might be called "reliability" of a device. This value, it is hoped, coincides more or less with what seems intuitively the appropriate one. However, none of the examples suggests what could be done about reliability if its value turns out lower than desirable. It will be the purpose of this section, and of the next one, to illustrate how the reliability concept introduced here can be used to study possible improvements. This section in particular will be concerned with the idea of preventive maintenance.

The general problem of when to attempt preventive maintenance, and of how to design an optimal one for a given operation, is by all indications a difficult one. In certain special cases, however, solutions can be obtained. One such case will now be discussed, as an illustration of the whole concept.

Consider a device such as the radio tube mentioned earlier (example 2 in the preceding section) which can be thought of as being replaceable an unlimited number of times. However, instead of using it until it fails, as was assumed in that example, a preventive replacement at a suitably chosen age T is to be considered. The economics of the situation is this: The price of the tube is C, and that must be paid in any case whether the replacement is preventive or a repair. In the latter case, however, an additional service charge P is expected. The replacement age T is to be so chosen that the mean rate of expenditure (the unreliability in this case) is as low as possible.

Example 2 has yielded a result which applies here: If no preventive maintenance is carried out, the mean rate of expenditure will be equal to the mean number of failures per unit time (which is 1/m,

in the steady state), multiplied with the cost of each failure (which is C+P). That is to say, the reliability in this case is

$$R = (C+P)/m$$

A study of the question of preventive maintenance yields the following facts. There are certain circumstances under which preventive maintenance is definitely uneconomical, in the sense that it can only increase expenditures. This is the case where the failure rate is either constant (which implies the exponential law), or uniformly decreasing. In the opposite case, preventive maintenance may, but need not, be profitable. If it is, and if T is chosen in an optimal fashion, the expenditures can be reduced from the value of R shown above, to

$$R^1 = P\, g(T)$$

In the case of the normal failure law, the optimal replacement age can be read from Fig. 3. It is assumed that one knows (or can make a good guess of) the mean life m and the σ of the distribution, and that the cost C and service charge P can be determined. The chart is then entered with the appropriate ratios and the correct replacement age, in multiples of σ, is immediately evident. The cost of the operation then follows from the failure rate chart in Fig. 1.

Take, for instance, a component which costs $20. It obeys the normal law with a mean life of 2 years, and a σ of 3 months. The service charge for the repair of a failure is $10. In this case, $m/\sigma = 8$, $C/P = 2$, and the chart of Fig. 3 yields $T/\sigma = 7$. This is to say, the device should be preventively replaced at the age of 21 months. The mean rate of expenditures (the "reliability" in this case) would be $1.25 per month without preventive maintenance, and 96¢ per month with it. This is a saving of nearly 25 per cent.

Fig. 3 exhibits a phenomenon which may be worth pointing out: The curves are missing in the left upper region, that is, for small values of m/σ and for large values of C/P. This is no accidental omission. The empty area indicates that no preventive maintenance should be done in that region.

5. Duplication of components

The duplication of components is a measure which, like preventive maintenance, can be adopted to improve the reliability of a device. The point will be to show by a brief example that the reliability concept introduced here can be used to evaluate the extent of such an improvement.

Consider again the situation treated in example 2, that is, a replaceable component such as a radio tube. To be specific, imagine this tube to be part of a broadcast transmitter. In such an installation, a duplicate transmitter is usually provided which is put on the air when the original one fails, to gain time for its repair. The question is, how much of an improvement in reliability is derived from that?

In example 2, a unit loss was charged for each failure and the replacement rate was obtained as the appropriate measure of reliability.

In the present example, a unit loss will be charged only in case both transmitters fail, that is, when the duplicate went out, before the original could be repaired. To simplify things, we will assume the exponential law not only for the failures of the tube but also for the repair times.

With this assumption, the steady-state loss rate (that is, the mean number of double failures per unit time) is found to be

$$R = \frac{1}{m} \cdot \frac{1}{2 + \frac{m}{m_R}}$$

Here, m_R is the mean repair time. The factor $1/(2 + \frac{m}{m_R})$ is the improvement factor. It is the greater, the greater m is relative to m_R. Thus, if the mean life is ten times greater than the mean repair time, the reliability is improved by a factor of 12. It can be seen from this that a substantial gain is possible in this type of operation from use of duplicate equipment.

The operation described in this example is very favorably influenced by equipment duplication. Caution is indicated, however. The gain need not be as great in all cases and in some operations the question of whether it is, or is not, economical to duplicate is not clear cut. A careful analysis of what improvement can be expected will usually be very much in order.

Reference:

(1) A. J. Lotka, A contribution to the theory of self-renewing aggregates, with special reference to industrial replacement, Ann. Math. Stat.; Vol. 10 (1939), p. 1

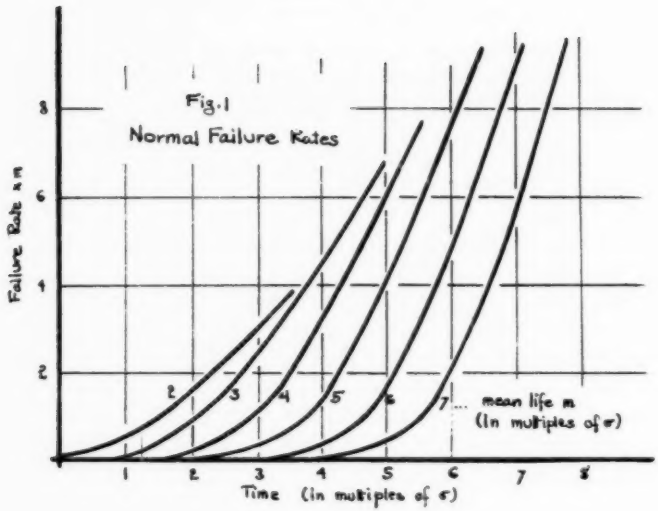


Fig. 1
Normal Failure Rates
Failure Rate x m
8
6
4
2
2
3
4
5
6
7
... mean life m
(in multiples of σ)
1
2
3
4
5
6
7
8
Time (in multiples of σ)

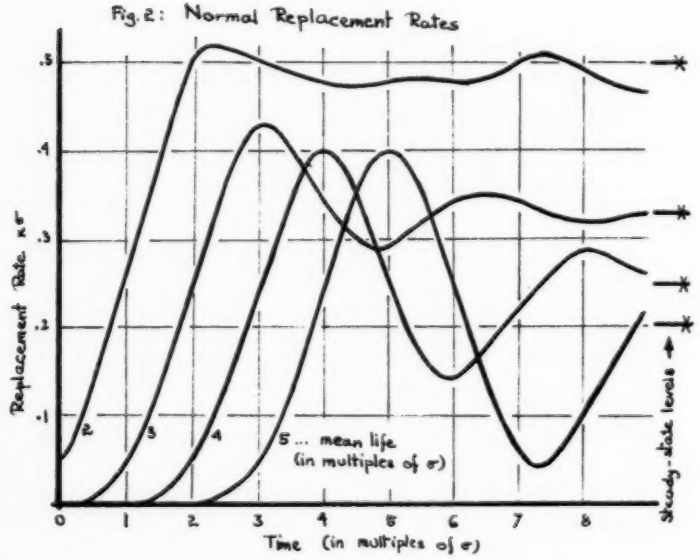
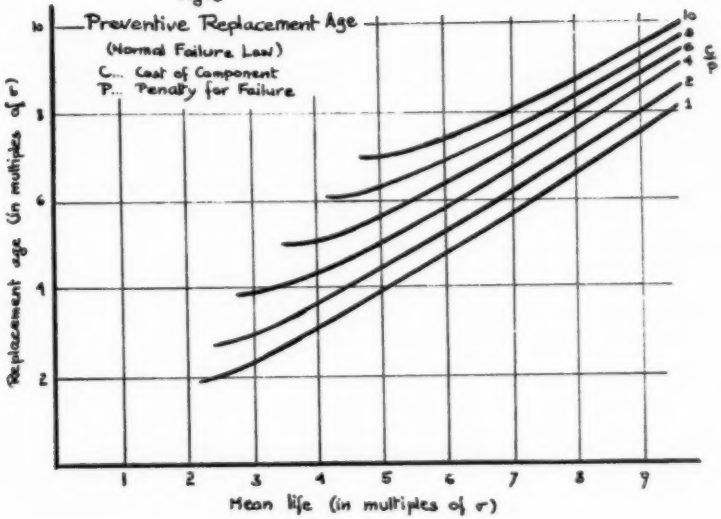


Fig. 2: Normal Replacement Rates
Replacement Rate x σ
.5
.4
.3
.2
.1
2
3
4
5 ... mean life
(in multiples of σ)
Steady-state levels
0
1
2
3
4
5
6
7
8
Time (in multiples of σ)

Fig. 3


Preventive Replacement Age
(Normal Failure Law)
C... Cost of Component
P... Penalty for Failure
Replacement age (in multiples of σ)
Mean life (in multiples of σ)
C/P
10
8
6
4
2
1

QUALITY REQUIREMENTS OF AUTOMOTIVE FABRICS

W. Ralph Strother
Ford Division of Ford Motor Company

Not too many years ago a car buyer had a choice of three or four body paint colors and the same number of fabrics. The fabrics were plain and durable and did not necessarily match the metal trim. Times have changed, the interiors and exteriors of our cars are bright, lively and as colorful as modern home interiors. The car buyer may choose practically any color in the rainbow with a matching interior. In addition to the wide choice of colors, there are many fabrics and tailoring schemes from which to choose.

In order for the stylists and engineers to accomplish this, many new fabrics, designs and fibers have been introduced into the automotive trim field. This rapid change to complex trim schemes has necessitated that attention be given to the development of detailed quality requirements which must be met by the fabric manufacturers.

We could probably speak indefinitely on those quality requirements which are of vital concern to four basic groups of people:

1. The Stylists, Designers and Engineers
2. The Trim Fabricators
3. The Plants that trim the bodies
4. Last, and most important, the Consumer

We will first discuss the overall quality requirements of automotive fabrics and then the following specific applications in the automobile body.

A. Headlinings
B. Seat Cushion and Back Covers
C. Convertible Tops
D. Carpets

General:

1. Quality Requirements of Automotive Fabrics

During the review and selection of fabrics by Styling and Engineering, the following basic quality requirements must be kept in mind:

Can the particular fabric be tailored to fit the body - such as headlinings, cushion covers, carpets, etc. without undue hardships to the fabricator?

Will the Fabric wear adequately: Seam strength, snag resistance, abrasion, etc.?

Will its colorfastness to light, crocking, bleeding, migration,

and atmospheric gases be satisfactory?

Does the Fabric have the proper hand, that is, feel?

Will the fabric sag, stretch, shrink or elongate too much and become undesirable to the customer?

Can the quality of the fabric be reproduced consistently from one shipment to another?

The difficulty of achieving all these necessary requirements has been increasing since World War II.

2. Quality Requirements of Fabrics for Specific Application

A. Headlining

The headlining is the covering of the roof interior. For many years the basic headlining materials were a napped woolen, wool blend, cotton, artificial leather and leather. However, changes have been made to conform with other styling changes in the car interior. Today you will find headlinings made from cotton, wool, synthetic fabrics, vinyl coated fabrics, woven paper, moulded plastic, masonite, cardboard and leather.

The more important requirements of a good fabric for headlinings are as follows:

1. Softness and Pliability - The material must span numerous radii without wrinkles.

2. Strength - The material must have sufficient strength to allow the assembly plant operator to exert the required amount of tension on the fabric during installation in order to have a wrinkle-free headlining without sags, wavy seams, etc.

3. Shrinkage - Shrinkage plays an important part in both the headlining final assembly and its life in the vehicle. No doubt, you have seen cars wherein the headlinings have pulled away from the back window area or from above the windshield. While there are numerous causes, one of the most important is excessive shrinkage in the material. On the other hand, some shrinkage is desirable, particularly in the assembly operation. Many minor wrinkles, sags, etc. may be removed by applying steam to the headlining after installation. Do not construe this to mean that steam is a cure-all. It is only an aid in correcting normal manufacturing variations.

4. Color Fastness - Color fastness to light is one of the most important quality requirements in headlining materials. Under modern styling trends the headlining is exposed to more light than ever before.

A crocking requirement was not too important in the past, since the headlining seldom came in contact with wearing apparel. In recent years, car roofs have been lowered and, in turn, the headlining is more likely to come in contact with hats. Therefore, the dye in the material should not

rub off easily.

5. Stretch and Set - How much will the material stretch under a constant load and what is the percentage of stretch that does not recover. If you have a material with high stretch and high set characteristics you will probably experience trouble in sagging and fullness in the end product either during assembly or in customer use. If a material varies widely in stretch characteristics, from one shipment to another, the headlining will vary from too small to too large as the master patterns are made and developed for the average stretch.

B. Body Cloth

Generally speaking, body cloth is the fabric used in the seat cushion and seat back covers.

Some of the more important requirements of body cloth are as follows:

1. Wear

When you are shopping for a suit or coat you ask yourself the question, will it wear well. In the automotive field we do the same thing. The amount of wear depends on many factors. Some very expensive fabrics have poor wearing qualities, yet have beautiful hand and appearance. However, duck is an inexpensive fabric which will wear well, but it is usually stiff and plain. Therefore, a compromise must be reached with the best possible wear for the use for which it is intended, yet have all of the other necessary characteristics that are required. Resistance to abrasion and snagging are two of the important wearing characteristics in body cloth.

2. Strength

Body cloth, with the exception of convertible top material, requires the greatest strength of all fabrics used in the car. It is constantly under stress and strain when the seat is occupied. There is greater strain applied to this fabric in the assembling of the seat covers to the springs, than in any other installation. If the material is cut undersize or shrinks slightly after it is sewn into a cushion or back cover, the operators assembling the cover will exert great strain on the fabric. Here is where stretch is required of the fabric.

3. Resistance to Fraying or Ravelling of Raw Edges

Normally, raw seam edges in automotive fabrics are not bound or pinked to prevent ravelling of the yarn. A fabric that ravels or frays easily may cause trouble in the completed cushion cover in service; the seams will separate due to loss of seam width. In some instances the material is edge-folded to prevent ravelling or yarn slippage; however this is a costly operation and may be avoided in most cases by choosing a fabric with a tight weave and applying a bonding agent to the back of the fabric.

4. Stretch and Set

The lack of adequate stretch and set characteristics of body fabrics probably represent one of the most serious problems that the automotive industry is encountering today. The ideal fabric is one that will stretch and completely recover to produce a wrinkle-free seat when it is in the showroom and still be in the same condition months later in service.

5. Elongation Resistance

The characteristic of a fabric whereby it grows or gets larger is usually due to increase in moisture content. This is quite common in all rayon fabrics. Fabrics should have very little or no elongation.

6. Color Fastness

Color fastness to light, wet and dry crocking, water staining and perspiration staining are quite important on all body fabrics which are constantly exposed to light and in contact with the passenger's clothing and 1001 other articles which are transported in a car.

7. Shrinkage Resistance

Shrinkage due to moisture in woven body fabrics is one of the less important characteristics, however, it should not be overlooked.

C. Convertible Top Fabric

Material usually specified for convertible tops is normally a 2 or 3 ply fabric with a rubber compound between the plys.

This material is probably subjected to more abuse than any other fabric on the car and yet, is of extreme importance, not only in its functional role but in the appearance of the car. Therefore, it is of the utmost importance to have a strong material that wears well, is dimensionally stable, (the material does not have excessive shrinkage or growth characteristics), and of superior color fastness to light and varying atmospheric conditions.

Some of the important quality requirements of Convertible Top Fabrics are as follows:

1. Strength - The material must be strong enough to withstand the constant strain exerted by air pressure when driving.

2. Color Fastness - To withstand the sun and weather exposure to which it is subjected.

3. Flexibility - The material should be flexible enough to fold without making too large a package or wrinkling excessively when in the stored position.

4. Shrinkage Resistance - Shrinkage has probably caused more customer dissatisfaction than all other characteristics

combined. The top becomes harder to operate, the top frame along the sides is exposed, and there are possibilities of water leaks.

5. Stretch and Set - As in most all automotive fabrics stretch and set are quite important in Convertible Top Fabrics - particularly in the assembling of the fabric deck to the frame. The operator must apply tension to the fabric in securing the fabric deck to the frame to remove wrinkles and any fullness that might exist.

6. Adhesion - The strength required to separate the outer fabric from the inner fabric must be adequate and uniform throughout the material.

7. Resistance to Wicking - The fabric should not wick when normally wet out. Wicking usually occurs around the seams where moisture comes in contact with raw edges of the fabric. Oftentimes the lining fabric of a top will become wet on a fabric that wicks badly.

8. Resistance to Elongation - A characteristic of a fabric to grow or expand is usually due to moisture absorption. This condition is generally prevalent in rayon top materials and can cause considerable trouble in the end product if not controlled closely, resulting in sags and wrinkles.

D. Carpets

Carpets in the car of today are quite similar to those in our home, they too have taken on the modern look. You will find carpets in all shades and of all types of fibers.

Some of the more important quality requirements of carpets are as follows:

1. Wear - Once again we have a fabric that is subject to considerable wear, but the areas that are subject to the greatest wear are usually protected by an additional reinforcement such as rubber foot or heel rests.

2. Shrinkage Resistance - Shrinkage due to excessive moisture may create problems after the carpets have been in use, in that the carpet pulls away from the scuff plates (metal covering over the door sill), resulting in the floor pan being exposed and the carpet edges rolling up.

3. Snag Resistance - Snagging or yarns pulling away from the fabric backing is most commonly associated with loop pile carpeting. If the pile pulls away from the backing easily, trouble may result in the cutting and sewing operation, assembly to the body, and in service.

The amount of research and development required on fabrics by the automotive manufacturers and textile mills has expanded tremendously and will probably increase as new styles, fabrics and fibers are introduced. Frequently, when a new fabric is introduced, it presents new problems in that a certain characteristic has changed slightly or

certain styling and fabrication changes require a characteristic to become more important than in the past. Therefore, you can see that the task of evaluating and revising quality requirements is a continuous job.

The Ford Motor Company has been, and still is, conducting an extensive research and development program on automotive fabrics to assure the consumer of quality in the new materials that are constantly being introduced.

One approach in our quality research and development program is that of working jointly with the fabric suppliers on particular problems to be solved. For example, a new convertible top fabric was introduced four or five years ago. The material had superior colorfastness to light, weather, age, etc., which was not available in other comparable fabrics, but the fabric had a high shrinkage factor. By working closely with the textile mills and fabric finishers the shrinkage was reduced below that of the previously specified material. As the result of the solution of the shrinkage problem in the top material, Ford has been in a position to offer a black convertible top whose color would remain virtually unchanged throughout its life.

Another example is the improvement on woven plastic ("Saran") body fabric as a partial result of the Ford-Dow technical meetings. Some of the improvements were light stability, shrinkage resistance, reduction of static electricity, improved color range, etc.

In summation, the Automotive Industries and the fabric suppliers are primarily interested in assuring that the fabrics will be satisfactory to the consumer. To fulfill our obligation to the consumer, it is of utmost importance that adequate control be maintained on the quality requirements we have just discussed, and on the other quality requisites not specifically presented in this paper.

QUALITY CONTROL'S OBLIGATION TO MANAGEMENT AND CUSTOMER ON RELIABILITY

Colonel Clair A. Peterson

I have been asked to present my views on the subject of Quality Control's obligation to management and the customer on reliability of complex equipment.

To state the case in simple brief terms, it is the obligation of Quality Control to assure management that the products for which management is contractually committed to furnish to its customer are in fact in accordance with the specified requirements.

This job of having to assure quality and reliability of the product places the quality control manager in the position of having to be concerned with all the functions and activities of the company which, in one way or another, have a bearing on the quality of the final product.

Furthermore, with the introduction of specific reliability requirements, the function of quality control becomes more widespread than heretofore. It is now necessary that the quality manager be directly concerned with the proving-out phase, as well as the in-service usage of the equipment.

This necessitates the establishment of a working relationship with the customer, so that operational performance can be closely followed and quality control problems quickly resolved.

Because of the important function of quality control in a reliability program, it is also necessary that the quality control manager participate in the negotiation and proposal process of contracting; so that all pertinent aspects of the contract and specifications are properly considered and provided for in the proposal.

Likewise, it is necessary that the quality control manager participate in the program for subcontracting and source selection. It is also necessary that he maintain, at least during the initial stage of production, an active liaison with those suppliers furnishing the more important items of equipment.

In the brief period of time allotted to me, it is not possible to cover all aspects of the quality control function. I can merely highlight some of the particular activities of quality control in carrying out its obligation to management and the customer.

We know that as we progress further into the concept of the weapon system, the role of the quality control manager becomes more and more diversified - extending his interests into outlying areas, and geographic locations remote from the manufacturing facility.

Under these new concepts, both the contractor and quality control continue to have certain obligations to the customer - even after the equipment has been accepted and placed in service.

"EVALUATING THE VENDOR"

W.A. MacCrehan, Jr.
Bendix Radio Division
Bendix Aviation Corporation

Vendor Quality Evaluation operates in four basic areas in which management activity, through Quality Control, is yielding dollar dividends and customer assurance.

The first is a basic understanding of the concept of variances. All products will not be alike. The degree of allowable difference that exists between one unit of a production run and another unit from the production run is controlled by the engineer through permissable variations, called specification tolerances. The desired dimension, or value, or characteristic of the design is called the nominal. Because of normal variances, as well as assignable cause of variances, the nominal characteristic will not be present in each unit of production. Consequently, the control of this variance to parameters within the specification tolerances is the heart of good Receiving Inspection ... therefore, raw material, and components controlled to insure that the design intent of product variances is being maintained, is a beginning of dimensional control of the vendor's product. At this point, to evaluate performance of the vendor it is necessary to establish some sort of "box score", a scale from zero to one hundred ... a direct listing of the estimated percentage defective found in sampling the product when received or a weighted characteristic ranking method; all are practical operations to evaluate the ability of the vendor to perform within the requirement of the specification.

A second basic concept is an understanding that three factors operate when a product is purchased ... price, delivery, and quality. Depending on the market availability, depending on the supply and demand of the particular product, and depending on competition among vendors, these three characteristics will vary in their importance. However, it is not good management to purchase on the basis of purely price and delivery, and a prime point of evaluating the vendor is to secure the whole-hearted cooperation of Purchasing in considering the factor of quality on an equal level with price and delivery when making purchasing considerations.

A third important concept is the knowledge of the error of sampling. Small samples taken from large lots yield a minimum of assurance. The error is not in the statistical sampling theory ... it lies in the lack of knowledge that the number of samples on which one bases an estimate is proportional to the assurance you desire that the decision you have reached is valid. Past history has indicated many case examples of successful non-statistical "ten percent of the lot" sampling plans that work. However, for every instance on record where "rule of thumb" non-statistical guessing is used, there are many more which show such a fixed plan is not good management because of the instability of the risk taken on insecure knowledge of the true ratio of good pieces to bad in the entire lot. Therefore, the sliding scale that ratios sample size to lot size used by modern sampling tables, such as Dodge-Romig, Mil Std. 105A, or various sequential sampling plans, is based on a known desire of confidence in the estimated percentage defective obtained through the ratio of good pieces to bad in the sampled lot. Further progressive thinking, such as continuous sampling plans used by Bell Laboratories and Navy Ordnance, utilizes this knowledge of sampling variances that the estimated

percentage defective is accurate within certain confidence parameters as justification to reduce the cost of inspection and test at Incoming Inspection and on the production line.

A fourth vital concept is a basic belief that all vendors would like to do a good job of supplying their customers ... and with the belief in this axiom, a desire by the customer to assist the vendor through an education program of histograms, simple running control charts and Lot Plot techniques invariably return good results of mutual benefit to both the vendor and the customer.

SUMMARY: All control is inherently based on the "cost of conformance versus the cost of control". To establish a quality evaluation program for vendor items takes an initial expense in these areas: people to make a preliminary survey of the vendor's facilities; people to audit the methods by which good or bad quality is determined at Receiving or Line Inspection; third, people to liaison between the vendor, the Purchasing Department and individuals concerned with the vendor's product quality on the assembly lines; and people to evaluate and instigate corrective action where your evaluation indicates unsatisfactory performance. Depending on the size of the organization, this could be a single person doing a multiple of these duties, or a specialist of proper technical background to cover each. The experience of many companies that operate an agressive vendor rating program has consistently shown it to be good management to invest quality dollars in this area of activity.

Control of variances present in a process can be quickly summarized into two categories: those variances that are known, and those variances that are unknown. The conditions causing known variances are usually corrected rapidly and with little investigation. Where variances are unknown as to their cause, statistical quality control methods through trend charts, process capability charts, Lot Plots, and similar techniques, spotlight the offending causes of variances with a minimum of time and effort.

Evaluating the vendor is a method by which you can separate the knowns from the unknowns and instigate corrective action to whatever degree seems to be necessary, based on the managerial evaluation of the weight of the problem in terms of delays to production, excess scrap and rework, and obsolescence or deterioration through excess inventory. All of these problems are faced daily and most firms either use a high volume of raw material that they complete into finished goods or buy multiple components for assembly into finished products. Therefore, a certain percentage of all end product quality is in the hands of your suppliers.. .. it is good management to control this element of your end product quality.

QUALITY CONTROL OF VISUAL CHARACTERISTICS

C. W. Carter
Rath & Strong, Inc.
Boston, Massachusetts

The most widely used of all sensing instruments is the human eye. By definition, then, a visual characteristic is any that may be sensed by eyesight. This would embrace all measuring instruments or gage readings, and here a distinction must be made.

That distinction will be between "dimensional" control and the control of visual characteristics. Considering dimensional control in its broadest sense, then, one should include all characteristics which may be gaged with some physical instrument and whose results are readily expressed in numerical terms. Examples of testing instruments for the control of dimensional characteristics range from a weighing scale to a temperature thermometer; from a micrometer to an ammeter; from a hardness tester to a pH indicator, and so forth.

Industries in which the control of quality is a matter of meeting a dimensional requirement are many. It is in this area that statistical quality control made its start and has effected its largest contribution. This would include attribute inspection, for therein the results are readily expressed in numbers, i.e. good = 0, bad = 1; and the percent defective is the measure of quality.

Visual characteristics, by contrast, may be considered as those appearance considerations which make the difference between satisfactory and unsatisfactory product. These may consist of specific defects, such as scratches, or dents; they may comprise the overall "looks" of a product, such as a fabric, or a television set. In either case, the quality could hardly be measured by conventional means. For example, it would be difficult to imagine an instrument capable of completely describing the quality of a scratch.

The control problem usually arises from manufacturing the product to dimensional specifications while selling it on the basis of visual characteristics. Experience indicates that this is a fairly universal problem; equally - and in some cases more - important than dimensional control. It is our opinion that much more attention should be given to this important phase of quality control.

It is the purpose of this paper to suggest that practical as well as statistically sound approaches can be successfully utilized to control the appearance quality of a product.

The general approach to any control problem may logically follow these four steps:

1. Determine the process capabilities
2. Eliminate or reduce principal sources of variation
3. Design an adequate control routine
4. Provide proper maintenance.

This paper will touch principally upon the first of these since it is probably the most difficult.

In order to define the process capabilities, one must be able to measure the quality characteristic under consideration. What sets this subject apart from the usual control problem is that herein we are presented with unusual measurement challenges.

Let us consider four specific solutions to measurement problems involving visual characteristics:

First, it may be possible to devise an instrument which will measure the particular characteristic. One such example is the profilometer, which measures metallic finishes by averaging the surface roughness. Another example is the spectrophotometer, measuring color and expressing it in terms of tri-stimulus values --- and so on.

Secondly, product specimens have been established as visual quality standards. These serve as inspection guides to the acceptable limit of appearance quality. This may then be handled on an attribute basis, and the quality level expressed as a percent defective. Where it is infeasible to transport the actual samples, 3-D pictures are sometimes employed. For instance, the Quartermaster Corps developed such a booklet showing allowable and not-allowable examples of a rather complete line of fabric defects.

The third solution is useful in instances where it is possible to isolate and study a particular visual defect by itself. This starts with the definition of a few innate descriptive categories within which the defect may be fully classified.

For example, it was desired to reduce the occurrence of dents in the manufacture of coffee percolators. The four descriptive categories selected were:

1. Location
2. Size
3. Source
4. Style.

Figure 1 shows the numerical assignments made to these four categories. By use of these numbers, then, any dent could be adequately described. Since each category was independent of the others, a total of 8x3x2x2 or 96 numerical combinations were possible; 1111, 1112, 1121, 1122, and so on through 8322.

If the dents were caused by purely chance reasons, each combination or kind of dent, would have an equal likelihood of occurrence. Out of, say, 500 dents observed there would be about 4-6 of each kind. However, if the distribution of dents were significantly different from the expected, - if there were any pronounced concentrations - some systematic assignable cause would be at work. Further, the nature of this cause might be indicated by the particular kind of frequently-occurring dent.

In order to get this picture, as well as the occurrence by operation a random-time-sampling was set up after the seven principal operations. Five percolators were carefully examined on each visit, and any observed dents were classified. The Audit Classification Work Sheet used in this study is shown in Figure 2. Following a period of approximately one month the classification data was summarized for the various sampling points.

The results appear in Figure 3. From this information, it was possible to detail the relative contribution of the several areas.

This is illustrated graphically in Figure 4.

The last solution is the utilization of opinion-ratings as a means of describing the overall visual quality of the product. This was suggested by two successful applications of subjective judgments in discerning differences in taste. One of these was at Wilbur-Suchard in the control of chocolate quality; the other at the Schenley Distilleries in the blending of whiskey. The plan developed and practiced at the Bigelow-Sanford Carpet Company [1] is an excellent illustration of this technique.

A more recent example is its application to the controlling of Acoustical Tile quality.

Here, three principal appearance categories were established:

	Category	Operation
1.	Machine	Base Board Machine
2.	Coating	Coaters
3.	Fabrication	Finishing (inc. cutting, drilling, bevelling, etc.)

As shown, each category relates directly to a particular operation in the manufacturing process. Each is relatively independent of the other. Thus, each category may be rated, according to an Inspector's judgment, practically without reference to the other two.

The point at which this inspection is done is the final sorting tables. Random-time-schedules, based upon the 96 5-minute periods in an 8-hour workday, were developed for the Inspector. At each sorting station, he samples five tiles each from the cull pile and the accepted product stack. Each tile is then rated in each of the three categories, using the following schedule of rating numbers:

Rating Number	Quality Description
0	Ideal
1	Incidental
2	Minor
3	Major
4	Critical

The sum of the three ratings gives the total quality rating for that tile.

Following an introductory, or familiarization period, a base period of one month served to establish (a) the process capability, and (b) the consistency of the Inspector's rating. During this period, samples were taken of the production coming to the sorter as well as the culls and good following sorting. Although different samples were selected for the "before" and "after" ratings, it was assumed that they came from the same population. The accepted tile ratings and the culled tile ratings were combined in the same ratio as the overall percent perfect and culls from Production records. This gave a prediction of the quality rating of tile coming to the sorter, which was then compared with the actual.[2]

Figure 5 shows this comparison. Since these were sufficiently similar, the "same population" assumption could be justified, and the consistency of the Inspector's rating was established.

This distribution of total quality ratings provided the gage as to the process capability for the base period. The average of the highest rating for each tile gives the quality level for that period. Thus the visual quality for the operation may be compared, as time passes, to show changes in the process.

A representative "control panel" of 100 tiles was selected and therewith the Inspector periodically "calibrates" his judgment. Further, this control was submitted to several customers for their rating. With two startling exceptions (in one case, the customer was more severe, in the other, less!), the results were very comparable. The Inspector's ratings were then adjusted to reflect these differences.

Another phase is evaluating the relative contribution of the several defects making up the quality problem. With each rating of 2 or greater, the Inspector must list the defect involved. The frequencies of each defect are then tabulated, and this serves as a continuing guide to production personnel for current and effective defect prevention.

Using the highest rating for each tile, considering 3 and 4 to be imperfect, the quality balance for the sorting operation may then be determined. Initially, this appeared as pictured in Figure 6.

That this control has made a positive contribution to quality improvement may be seen from the following table:

Item	Base Period	After Six Months
Percent Culls to Sorters	14 %	12 %
Percent Culls to Inventory	5 %	3 %
Percent Good in Culls	24 %	8 %
Average Rating	1.7	1.5

Perhaps the most immediate contribution was in the area of reducing the occurrence of good product in the cull pile. This was not accomplished at the expense of the outgoing quality level, and there is definite indication of an improvement in the visual quality of tile submitted to the sorters.

In conclusion, then, the quality control of visual characteristics may be accomplished in at least four manners:

1. Devise a measuring instrument
2. Establish comparison standards
3. Isolate a particular visual defect and describe it in quantitative terms
4. Utilize opinions as a means of rating the overall product quality.

Among these proven techniques, there should be one or a combination thereof to which your visual quality problem will yield. Only your application will really demonstrate how effective they can be.

(1) "Quality Control of Visual Characteristics", by Stanley E. Lezinski, I.Q.C., July 1953.

(2) These were compared using the X^2 technique.

PERCOLATOR DENT STUDY

Instruction Sheet

For the Classification of Dents

Without tracing its reason, one can numerically classify a dent according to four characteristics:

(1) Location on the Percolator
(2) Size of the dent
(3) Source
(4) Style

In order to conveniently analyze these data, the following code is used:

1. LOCATION:

Looking at the percolator from on top:

It is then divided into two horizontal sections and four cross-sections; thus, a dent located on the top half just to the right of the spout would be classified as #1 under the characteristic "Location".

2. SIZE OF DENT:

The longest dimension of the dent is the basis for classification of this characteristic:

SIZE OF LONGEST DIMENSION	CLASSIFICATION NUMBER
Under 1/4"	1
Between 1/4" and 1"	2
Over 1"	3

3. SOURCE:

	CLASSIFICATION NUMBER
"Outside - In"	1
"Inside - Out"	2

4. STYLE:

	CLASSIFICATION NUMBER
Sharp (either line or point)	1
Rolling	2

PERCOLATOR DENT STUDY

AUDIT CLASSIFICATION WORK SHEET

INSPECTION POINT_____ DATE_____

TIME	SEQUENCE NO.	LOCATION	SIZE	SOURCE	STYLE	NO. DENTS

PERCOLATOR DENT STUDY

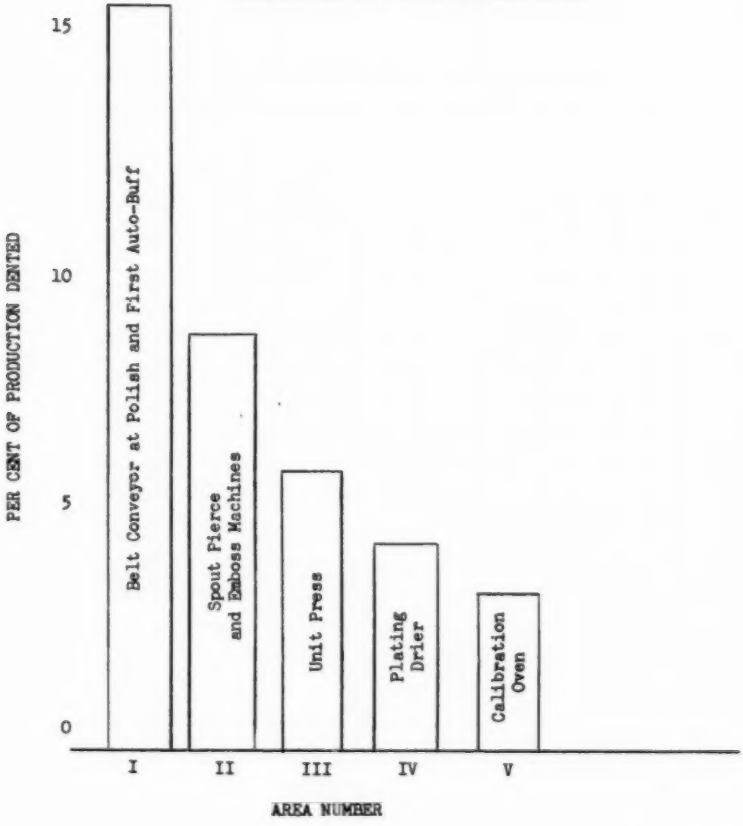
Summary of Dent Classifications at Seven Stations and the Five Special Lots Through Assembly

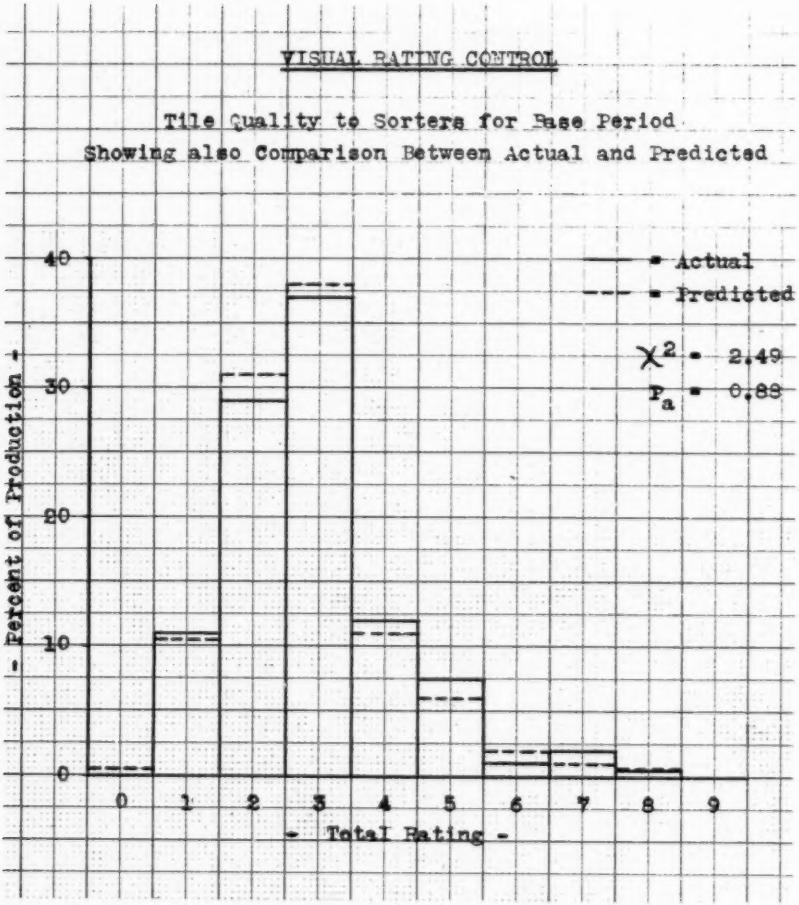
	After Polish		After Stud Weld		After Hand Color Strip		After Hand Color Rework		After Chrome Plating		Assembly		Dent Table		Five Special Lots	
	No. Dents	%	No. Dents	%	No. Dents	%	No. Dents	%	No. Dents	%	No. Dents	%	No. Dents	%	No. Dents	%
Total	42	100.0	95	100.0	77	100.0	38	100.0	83	100.0	84	100.0	302	100.0	75	100.0
Location 1	NA		0	0.0	2	2.6	0	0.0	1	1.2	1	1.2	5	1.7	1	1.3
2	NA		0	0.0	1	1.3	1	2.6	0	0.0	5	6.0	8	2.6	1	1.3
3	NA		1	1.1	0	0.0	0	0.0	3	3.6	5	6.0	16	5.3	3	4.0
4	NA		4	4.2	2	2.6	1	2.6	4	4.8	1	1.2	15	5.0	0	0.0
Top		0.0	5	5.3	5	6.5	2	5.2	8	9.6	12	14.4	44	14.6	5	6.6
5	NA		27	28.4	8	10.4	6	15.8	6	7.2	14	16.7	42	13.9	13	17.3
6	NA		20	21.0	18	23.3	5	13.2	13	15.7	14	16.7	45	14.9	13	17.3
7	NA		18	18.9	22	28.6	11	29.0	28	33.8	15	17.8	74	24.5	23	30.7
8	NA		25	26.4	24	31.2	14	36.8	28	33.8	29	34.4	97	32.1	21	28.1
Bottom	42	100.0	90	94.7	72	93.5	36	94.8	75	90.4	72	85.6	258	85.4	70	93.4
Size - Under 1/4"	25	59.5	42	44.2	46	59.7	25	65.8	65	78.3	65	77.4	187	61.9	44	58.7
1/4" - 1"	15	35.7	38	40.0	20	26.0	7	18.4	15	18.1	15	17.8	95	31.5	30	40.0
Over 1"	2	4.8	15	15.8	11	14.3	6	15.8	3	3.6	4	4.8	20	6.6	1	1.3
Source - Outside-In	42	100.0	80	84.2	55	71.4	30	79.0	31	37.4	36	42.9	138	45.7	35	46.6
Inside-Out	0	0.0	15	15.8	22	28.6	8	21.0	52	62.6	48	57.1	164	54.3	40	53.4
Style - Sharp	21	50.0	73	76.8	56	72.7	33	86.8	60	72.3	66	78.6	210	69.5	39	52.0
Rolling	21	50.0	22	23.2	21	27.3	5	13.2	23	27.7	18	21.4	92	30.5	36	48.0

NA - Not Applicable

PERCOLATOR DENT STUDY

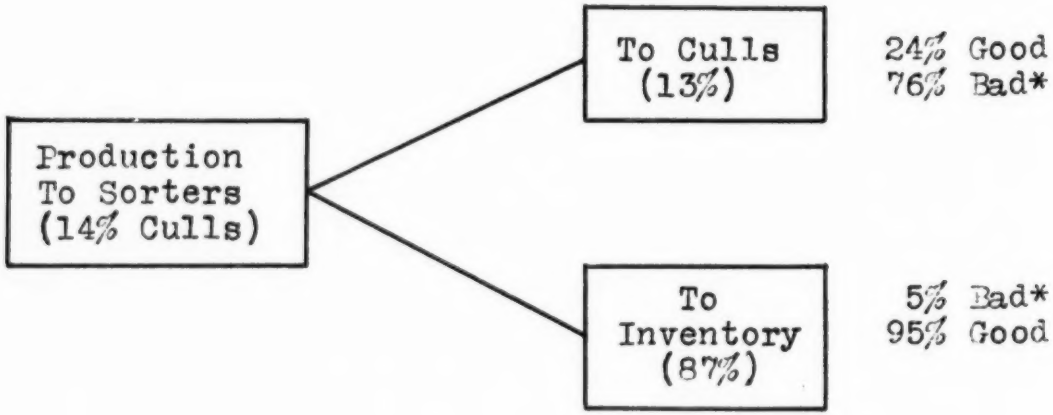
Graph Showing Relative
Contribution of Five Principal Areas




VISUAL RATING CONTROL
Tile Quality to Sorters for Base Period
Showing also Comparison Between Actual and Predicted
- Actual
--- - Predicted
χ^2 - 2.49
P_a - 0.88
- Percent of Production -
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9
- Total Rating -

VISUAL RATING CONTROL

Schematic Representation of Quality Levels
at Sorting Operation During Base Period



* Ratings of 3 or 4.

SELECTION AND RANKING PROBLEMS WITH BINOMIAL POPULATIONS

Shanti S. Gupta, Marilyn J. Huyett and Milton Sobel
Bell Telephone Laboratories

1. Introduction.

This paper deals with three problems concerned with the ranking of two or more binomial populations. In all the problems the goal is to find the best or the better populations; we define the best population to be the one with the highest probability p of a success on a single trial. In the first problem the goal is to find the number of observations required to guarantee with probability P* that the largest proportion of successes comes from the best population when a certain indifference-zone condition holds. In the second problem the goal is (for any given number of observations) to select a subset of the k given populations for which we can assert with probability P* that the best population is contained therein. We would like a procedure which makes the number of populations in the retained subset small, and at the same time large enough to satisfy the P*-probability condition given above. In the third problem the goal is (for any given number of observations) to select a subset of the k given populations for which we can assert with probability P* that all populations better than a given standard or control are contained therein.

The normal counterpart of problem I is treated by Bechhofer in [1]. The normal counterpart of problem II is treated by Gupta in [3]. A different treatment of the normal counterpart of problem III is given by Dunnett in [2]. Earlier, Paulson [6] considered for the normal and binomial cases the problem of selecting the best one of k categories (or populations) when one of the k populations is regarded as a standard or control. Paulson's and Dunnett's procedures both guarantee a probability P* of selecting the standard as best when all the populations have equal parameter values; in the procedures given below, a reasonable definition is given for a correct decision and a probability P* of a correct decision is guaranteed for all possible true configurations of the unknown parameter values. Applications are given in all of these papers. Any binomial or go-no go process in which it is desired to sift out the better processes or eliminate the poorer ones is a potential application of the problems treated in this paper. Many applications are not immediately apparent. For example, if units from several different manufacturing processes are put on a life test, the experimenter may wish to find the process which produces the most reliable unit. If he defines a unit to be good (i.e. a success) if it performs satisfactorily for a certain number of hours and to be poor (i.e. a failure) otherwise, then the binomial framework is applicable.

The corresponding problems for selecting the population with the smaller or smallest value of p are mathematically equivalent to the problems treated here and the same tables are usually applicable.

In problems II and III the formulation is more flexible in that the experimenter can select a subset and withhold judgment about which is the best one. Then the experimenter (or his boss) may wish to select one from this subset on the basis of economic or other considerations. It should be noted that in selecting a subset containing the best population we are automatically eliminating distinctly inferior populations so that these procedures may be regarded as elimination or screening

procedures.

2. Problem I.

Associated with each of the k given binomial populations there is an unknown probability p of success on a single trial. Let $\delta \geq 0$ denote the true (but unknown) difference between the largest p and the second largest p for the k given populations; for simplicity, we assume that $\delta > 0$. The goal of the experimenter is to select one of the given populations and assert that it has the largest p-value. If the population with the largest p-value is chosen then the selection is defined to be a correct selection. Before the experiment is performed the experimenter specifies two numbers (δ^*, P^*) with $\delta^* \geq 0$ and $\frac{1}{k} < P^* < 1$, and states that he requires a procedure which guarantees a probability P^* of a correct selection whenever the true difference δ is greater than or equal to δ^*.

The above requirement may be regarded as an indifference-zone condition since, roughly speaking, it states that the experimenter is willing to relax his probability requirement if the "error" is going to be one of selecting a population whose p-value differs from the largest p by less than the specified value δ^*.

Formally, we can define a selection as a correct decision if the selected population has a p-value which differs from the largest p by δ^* or less. Then the given procedure guarantees a probability P^* of a correct decision regardless of the true configuration of the unknown parameter values.

Procedure

The procedure will, of course, be to select as best the process which gives the highest proportion or number of successes; the only remaining problem is to determine the common number n of observations required per population to satisfy the indifference-zone condition. Table I gives the required number n for selected values of δ^* and P^*; a more complete set of tables with an explanation of their construction can be found in [7].

If there are ties for first place among the sample proportions of successes the experimenter may decide between contenders for first place by a chance experiment with equally probable alternatives. However, even if the experimenter wishes to use economic or other considerations to decide between contenders for first place, it appears reasonable in most problems to use Table I without any change. If the experimenter recommends all the contenders for first place and he is satisfied if one of them is the best one, then the values in Table I can be reduced by an amount equal to the largest integer in $(1/\delta^*)$.

A more complete write-up on this formulation of the problem together with another formulation which attempts to make use of à priori information is given in [7].

3. Problem II.

In this problem the number of observations is assumed to be given. Before or after the experiment is performed the experimenter specifies

one number P* with $\frac{1}{k}<P^*<1$ and states that he requires a procedure for selecting a subset of the k populations which guarantees with probability P* that the best population is contained in the selected subset. It is desirable to have a subset which is small and, at the same time, large enough to satisfy the above condition. It should be noted that there is no indifference-zone involved in the above requirement. We are assuming, for simplicity, that $\delta>0$ as in problem I so that the best population is uniquely defined.

If we define the selection of any subset containing the best population as a correct decision, then the proposed procedure guarantees a probability of a correct decision of at least P* regardless of the true configuration of the parameter values. If $\delta\geq 0$ and if we define the selection of a subset to be a correct decision if it contains at least one population with a largest p-value, then the proposed procedure still possesses the same property.

We consider two procedures for this problem; one is exact in the sense that it involves binomial theory, and the other is approximate in the sense that it employs an arc sine transformation and makes use of the normal approximation to the binomial. In the first case the number of observations is assumed to be the same for each population; in the second case this assumption is not made. The first procedure satisfies the requirement for all possible configurations but is more conservative and in most cases requires a subset as large or larger than that required by the second procedure; the second procedure has a probability of a correct decision which for certain small values of n and certain unfavorable configurations will dip below the specified value P* due to the inaccuracy of the normal approximation for small n. Even for the worst cases when there is a common p, this curve flattens out and approaches P* throughout the open interval, $0<p<1$, as n grows large.

The joint confidence statement that can be made after experimentation with confidence P* is that the parameter value of each eliminated population is less than that of the best population. It is shown in [3] that for the normal case other types of joint confidence statements can be made in place of that given above.

Procedure IIA

"Let x_i denote the number of successes observed in the n observations from the i^{th} population ($i = 1,2,\ldots,k$) and let x_{max} denote the maximum of these k integers. Retain the i^{th} population in the selected subset if x_i falls in the closed interval

(1) $$\left[x_{max} - d,\ x_{max}\right]$$

where d is given in Table IIA as a function of P* and n."

ProcedureIIB

"Make the arc sine transformation

(2) $$z_i = \sqrt{n_i + \frac{1}{2}}\left[\arcsin\sqrt{\frac{x_i}{n_i+1}} + \arcsin\sqrt{\frac{x_i+1}{n_i+1}}\right]$$

and let z_{max} denote the maximum of these numbers. Retain the i^{th} population in the selected subset if z_i falls in the closed interval

$$[z_{max} - d, z_{max}] \tag{3}$$

where d is given in Table IIB as a function of k and P* only."

For more complete forms of Table IIA and Table IIB the reader is referred to [5] and [3], respectively.

4. Problem III.

This problem is similar to problem II, the main difference being that one of the k populations is regarded as a standard or control and the others are compared with the standard. Before or after the experiment is performed the experimenter specifies one number P* with $1/2^{k-1} < P^* < 1$ and states that he requires a procedure for selecting a subset which guarantees with probability P* that all the populations as good as or better than the control are contained in the selected subset.

If we define the selection of any subset containing all populations as good as or better than the standard as a correct decision, then the proposed procedure guarantees a probability of a correct decision of at least P* regardless of the true configuration of the parameter values.

The joint confidence statement that can be made after experimentation with confidence P* is that the parameter value of each eliminated population is less than that of the standard.

We consider two cases according as the parameter value of the standard p_o is known or unknown.

Procedure IIIA (Known Standard)

"Let x_i denote the number of successes observed in n_i observations from the i^{th} population. Retain the i^{th} population if

$$\frac{x_i}{n_i} \geq p_o - d\sqrt{\frac{p_o(1-p_o)}{n_i}} \tag{4}$$

where d remains to be determined. If we let M_i denote the smallest integer greater than or equal to

$$m_i = n_i p_o - d\sqrt{n_i p_o (1-p_o)} \quad , \tag{5}$$

then the rule (4) can be written in the simpler form

$$x_i \geq M_i \quad " \tag{6}$$

For the special case $n_i = n$ $(i = 1,2,..,k-1)$ we have $M_i = M$ and the common value of M is given in Table IIIA for selected values of p_o, k,

P* and some small values of n. For large values of n_i (not necessarily equal) the value of d to be used with (4) can be conveniently obtained by the relation

$$F(d) = P^{*1/(k-1)} \tag{7}$$

where F(d) denotes the probability that a standard normal chance variable is less than d. Thus d is easily obtained with the help of any table of the integral of the normal distribution.

Procedure IIIB-1 (Unknown Standard)

"Retain the i^{th} population (i = 1,2,...,k-1) if the number x_i of observed successes in n_i observations satisfies the inequality

$$\frac{x_i}{n_i} \geq \frac{x_o}{n_o} - \frac{d'}{2}\sqrt{\frac{1}{n_i} + \frac{1}{n_o}} \tag{8}$$

where the subscript zero refers to the standard population and d' remains to be determined."

For the special case $n_i = n$ (i = 1,2,..,k-1) the rule (8) takes the simpler form

$$x_i \geq x_o - d \tag{9}$$

where d is given in Table IIA as a function of k, n and P*. In the more general case if the n_i are large and not necessarily equal, it is possible to use the following approximate procedure.

Procedure IIIB-2

"Let z_i be defined as in (2). Retain the i^{th} population if

$$z_i \geq z_o - d \tag{10}$$

where d is given in Table IIB as a function of k and P* only."

TABLE I

Minimum number of observations required per process to guarantee a probability P* of a correct selection when the true difference δ is at least δ^*. The four values in each cell correspond to the probability levels P* = .80, .90, .95, .99, respectively.

k \ d*	.05	.10	.15	.20	.30	.40	.50
2	142	36	16	9	4	3	2
	329	83	37	21	9	5	4
	541	135	60	34	15	9	5
	1082	270	120	67	29	16	10
3	273	69	31	17	8	5	3
	498	125	55	31	14	8	5
	735	184	82	46	20	11	7
	1308	327	145	81	35	20	12
4	359	90	40	23	10	6	4
	601	150	67	38	17	9	6
	850	212	94	53	23	13	8
	1442	360	160	89	39	21	13
10	606	151	67	38	17	10	6
	890	222	98	55	24	13	9
	1169	291	129	72	32	17	11
	1803	449	198	111	48	26	16

TABLE IIA

Values of d for Procedures IIA and IIIB-1 to be used where there is a common number n of observations from each of k populations. The two values in each cell correspond to the probability levels P* = .90 and .95, respectively.

n \ k	2	5	10	15	20	25	30	40	50
5	2 3	3 3	3 4	3 4	4 4	4 4	4 4	4 4	4 5
10	3 4	4 5	5 5	5 6	5 6	5 6	5 6	6 6	6 6
15	4 5	5 6	6 7	6 7	6 7	7 7	7 8	7 8	7 8
20	4 5	6 7	7 8	7 8	7 8	8 8	8 9	8 9	8 9
25	5 6	6 8	7 8	8 9	8 9	8 9	9 10	9 10	9 10
30	5 6	7 8	8 9	9 10	9 10	9 10	9 11	10 11	10 11
35	5 7	8 9	9 10	9 11	10 11	10 11	10 11	11 12	11 12
40	6 7	8 10	9 11	10 11	10 12	11 12	11 12	11 13	12 13
45	6 8	9 10	10 11	11 12	11 12	11 13	12 13	12 13	12 14
50	6 8	9 11	11 12	11 13	12 13	12 13	12 14	13 14	13 14

TABLE IIB

Values of d for Procedures IIB and IIIB-2 to be used when the numbers n_i of observations are large for each of the k populations.

P* \ k	2	5	10	15	20	25	30	40	50
.75	0.95	1.85	2.26	2.47	2.60	2.70	2.78	2.89	2.98
.90	1.81	2.60	2.98	3.17	3.30	3.39	3.46	3.58	3.66
.95	2.33	3.06	3.42	3.60	3.72	3.81	3.88	3.99	4.07
.99	3.29	3.92	4.25	4.41	4.52	4.61	4.67	4.77	4.85

TABLE IIIA

Smallest integer value M required for Procedure IIIA for selected values of n, p_o, k and P*. The three values in each cell correspond to k = 2, 3 and 5, respectively.

P* = .90 (Below double line, P* = .95)

n \ p_o	.10	.20	.30	.40	.50	.60	.70	.80	.90
5	0	0	0	1	1	2	2	3	4
	0	0	0	0	1	1	2	2	3
	0	0	0	0	0	1	1	2	3
10	0	0	1	2	3	4	5	6	8
	0	0	1	2	2	3	5	6	7
	0	0	0	1	2	3	4	5	7
15	0	1	2	4	5	7	8	10	12
	0	1	2	3	4	6	8	9	11
	0	0	1	2	4	5	7	9	11
20	0	2	3	5	7	9	11	14	16
	0	1	3	5	6	8	11	13	16
	0	1	2	4	6	8	10	12	15
25	1	3	5	7	9	12	15	17	21
	0	2	4	6	8	11	14	17	20
	0	1	3	5	8	10	13	16	19
50	2	6	11	16	20	26	31	36	42
	2	6	10	14	19	24	30	35	41
	1	5	9	13	18	23	29	34	41
5	0	0	0	0	1	1	2	2	3
	0	0	0	0	0	1	1	2	3
	0	0	0	0	0	1	1	2	3
10	0	0	1	2	2	3	5	6	7
	0	0	0	1	2	3	4	5	7
	0	0	0	1	2	3	4	5	6
15	0	1	2	3	4	6	7	9	11
	0	0	1	2	4	5	7	9	11
	0	0	1	2	3	5	6	8	10
20	0	1	3	4	6	8	11	13	16
	0	1	2	4	6	8	10	12	15
	0	1	2	3	5	7	9	12	15
25	0	2	4	6	8	11	14	17	20
	0	1	3	5	8	10	13	16	19
	0	1	3	5	7	9	12	15	19
50	2	6	10	14	19	24	30	35	41
	1	5	9	13	18	23	29	34	41
	1	4	8	12	17	22	28	33	40

REFERENCES

[1] Bechhofer, R. E., "A single-sample multiple decision procedure for ranking means of normal populations with known variance", Ann. Math. Stat. Vol. 25 (1954), pp. 16-29.

[2] Dunnett, C. W., "A multiple comparison procedure for comparing several treatments with a control", J. Amer. Stat. Assn., Vol. 50 (1955), pp. 1096-1121.

[3] Gupta, S. S., "On a decision rule for a problem in ranking means", Mimeograph Series No. 150, (May 1956), Institute of Statistics, Univ. of North Carolina, Chapel Hill, N. C.

[4] Gupta, S. S. and Sobel, M., "On selecting a subset which contains all populations better than a standard", submitted for publication in the Ann. Math. Stat.

[5] Gupta, S. S. and Sobel, M., "A procedure for selecting a subset containing the best of k binomial populations", to be published.

[6] Paulson, E., "On the comparison of several experimental categories with a control", Ann. Math. Stat., Vol. 23 (1952), pp. 239-246.

[7] Sobel, M. and Huyett, M. J., "Selecting the best one of several binomial populations", Bell System Tech. J., Vol. 36 (1957), pp. 537-576.

QUALITY CONTROL IN THE CARTON PLANT

D. R. Carlson
The Gardner Board and Carton Co.

What is Quality?

Quality is the conformance to functional and aesthetic standards for which a price has been agreed upon and contracted with the customer.

What is Quality Control?

Quality Control is the sum total of all methods used to attain quality.

What is Statistical Quality Control?

Statistical Quality Control is a systematic procedure under which random and representative raw data is gathered and analyzed to determine whether all or any part of such data is drawn from within or without a chance cause system. Such analysis to result in conclusions based on predetermined levels of confidence and probability. The intelligent interpretation of these conclusions aid the identification of the when, where and frequently how of the assignable causes resulting in departure from the desired chance cause system as well as a better understanding of process capabilities.

Why Quality Control at Gardner's?

1. To help make production of quality as easy as possible.
2. To help provide incentives to people to do good work.
3. To help protect the customers' interests in our plants.
4. To deliver production news - not history - and facts not opinions - to operating crews.
5. To supply facts to operating management by:
    a. Measuring the adherence to or deviation from specifications.
    b. Establishing procedures for rapid feed-back of information to operating people to aid the decision making process on which control is based.
    c. Analyzing whether measured deviations are those which can be considered due to chance causes or possible assignable causes which should be investigated, isolated, eliminated or made a permanent part of the process.
    d. Indicating the degree to which raw or semi-finished materials fail to meet specifications so that they can be segregated for special handling.
    e. Summarizing inspection data and reporting

periodically to operating management so that long range process trends can be detected.

6. To help determine the characteristics which reflect desired performance of the finished product.
7. To aid in the establishment of specifications.
8. To help establish test methods which can measure whether specifications are being met.
9. To help determine process capabilities.
10. To aid technical people in data analysis and design of experiments.

What Limits the Application of Quality Control Techniques?

Quality Control techniques can be applied under any circumstance where data can be gathered. Quality Control requires only that a quality characteristic, a condition, or a set of circumstances be identified and measurable, whether objectively or subjectively, to be subject to control procedures.

Philosophy of Quality Control at Gardner

Quality Control to function adequately and develop must have the support of all levels of management.

Quality Control is a staff department and functions in an advisory capacity. Quality Control does not directly stop machines, processes or reject material.

The direct responsibility for quality lies with Production.

The Quality Control Department shall be completely independent from production and function as an internal auditor of quality.

The Quality Control system is a friend of the workman and not a monitor over him.

Quality Control shall evaluate each and all samples as if they represented 100% of production. Failure to meet specifications must be reported without bias.

The basis on which finished product is shipped may not always be consistent. Normal shipping requirements may be affected by customer's immediate need for cartons, unusual operating problems, agreed upon price changes with customer, etc. Quality measurement standards, must, however, remain consistent. A constant reference base is necessary so that changes can be detected not only in the quality level but in the customer's attitude toward quality characteristics.

Quality Control in the Printing Plant

The problem of establishing a quality control system has two main factors. One is insuring that the techniques of measurement and analysis are both accurate and valid.

The other factor is the practical difficulty of selling a program which contains inherently some unpopular characteristics. In the following discussion, time and space permit only a few of the techniques which are used for quality control and analysis. Some of the industrial relations problems will also be discussed. Much has yet to be done to completely sell people on a program which they got along without for years. There is objection to a system which insists on publicizing (for all to see) a man's mistakes even though it also publicizes (for the same audience) his good work.

In-Process Inspection - Printing

A printed sheet is selected at random at least once an hour from each press. This sample sheet is taken to the inspectors' Quality Control station and placed on a slant top table with a standard light. The sheet is inspected carton by carton. The basic inspection unit is the carton. Defective areas on the cartons are circled and identified. Defects are classified into three categories of severity. To each of these categories is assigned demerits or severity points. These severity classifications are defined as:

Minor - Minor defects by definition will always be accepted by the customer. In our minds, however, it is a defect and should become part of the defective quality record. One severity point is assigned to this classification.

Major - Major defects are worse than a minor (just noticeable) defect, but are not serious enough to reject the carton. There should not be too many. To this classification is assigned three severity points.

Critical - Critical defects by definition cause the carton to be rejected. To this classification is assigned six severity points.

Despite these definitions, the cumulative effect of several lesser defects on a carton have an adverse affect beyond the explicit meaning of the definition. That is, three minor defects on the same carton will usually result in as much concern as the presence of a single major defect. Six minor defects or two major defects on the same carton will generate nearly as much action as a single rejectable critical defect.

By definition a given defect can occur only once on a carton and is evaluated with the severity of its cumulative effect.

Because the pressroom has traditionally been the domain of men, Quality Control at Gardner was first set up so that,

if necessary, the inspection function could have been carried out by a deaf mute. Fortunately, it has never been necessary to prove this. Information is communicated to the pressmen in four different ways.

1. The Daily Control Chart

The plotted point for each test is the average number of severity points per carton on the sheet. Also on the chart appears the type of defect, its severity and the die position on which it was found.

2. Signals

On the back of each control chart board there are a red hand, a yellow hand and a green ball. When a #6 (critical defect) has been found on the sheet, the red hand goes up. If no more than a #3 (major defect) was found the yellow hand goes up. If no more than minor (#1) defects are found, the green ball stays up. The presence of both the red hand and the green ball indicates to the production foreman that at the last inspection a critical defect was found by the inspector and that the pressman had taken corrective action.

The red hand alone indicates to the production foreman that a critical defect was found by the inspector and that the condition has not yet been corrected. Fixed or not, the colored hand stays up until the following test when the quality situation is revaluated.

3. Exhibit

The inspected sheet on which all defects are located and identified is returned to the press table.

4. Personal

The presence of a critical defect on a sample sheet is made known verbally and immediately to the pressmen.

There are a series of four integrated charts which use the same basic statistic (SP/c). These charts help us keep a finger on the pulse of quality variation from individual test to job average and from hourly fluctuation to yearly variation. Before we discuss these, however, there are a few factors which should be considered.

It is desired that the plotted points on the daily control chart reflect those things for which the pressman is responsible and accountable. However, there are times when the pressman is instructed to run with marginal materials or until the end of a feeder load or with a condition which cost factors indicate need not be completely eliminated. Defects occuring under those conditions are, therefore, not the direct responsibility of the pressman. Inspection and the quality record, however, go on. For this reason the foreman's okay section was added to the daily control chart.

When a situation like that described above occurs, the pressman goes to his foreman and gets an okay. The occurrence of that defect, thereafter, results in the defect being charged to the foreman's section of the chart. The graph line then continues to reflect those things for which the pressman is accountable.

Another important factor in in-process inspection is that each pressman is responsible only for his particular operation and is not penalized in the quality record because the material with which he is working may be relatively more valuable. Defective conditions from previous operations are not charged again during the current inspection unless these previous defects result in new defects. The plea that the operator or pressman is not responsible for the substandard material upon which he performs his function is only partly valid. Accountability there must be for the substandard material that adversely affects the quality record during any inspection. Its continued use must be the responsibility of Production and defects cause by it charged to the accountable source.

Figure 1 is a daily control chart. The figures entered in the body of the control chart are the die positions of the carton on which the listed defect was found. Die position numbers circled in blue are #3 defects, those circled in red are #6 defects, not circled are minor defects. The two colors are herein indicated by circles and rectangles respectively drawn around the die position number.

Control chart limits on repeat orders are based on standard values calculated from past data. On jobs for which no past data is available approximately one day's production (20-24 tests) is the basis for determining control chart values. When shifts in the process level occur, control limits are recalculated.

The value of the plotted point is determined by totaling all the severity points on the inspected sheet and dividing by the number of cartons on the sheet.

$$SP/c = \frac{TSP}{TCI} \begin{matrix} \text{(Total severity points)} \\ \text{(Total cartons inspected)} \end{matrix}$$

The process mean is $\overline{SP/c}$ and the standard deviation is $\sqrt{\frac{\overline{SP/c}}{k}}$ where k is the number of cartons on the sheet.

Control chart limits are given by $\overline{SP/c} \pm 3\sqrt{\frac{\overline{SP/c}}{k}}$, where $\overline{SP/c}$ is the process mean, $\sqrt{\frac{\overline{SP/c}}{k}}$ the standard deviation and k is the number of cartons up on the sheet.

The statistical validity of this approach is based on the:

a. Practical equivalence of the severity point ratings of the three severity classifications.
b. The limitation of defect occurrence to one per any one carton for any one given defect.
c. The comparable results of treating the same data as defects per carton and Severity Points Per Carton.

As a matter of fact, the control limits for SP/c are just a little bit tighter than D/c (defects per carton). The question might, therefore, be asked, "Why not use a defects per carton basis?" The answer is that operating personnel want to know not just that a defect is a defect but the type and severity of a defective condition to more intelligently initiate corrective action.

Another advantage of this approach has been that jobs running side by side can be compared even though they differ in size, copy, color and number of die positions on the sheet.

The daily control chart by itself, however, was inadequate for the following reasons. At 7:00 in the morning the daily control chart is changed and quality information from the past 24 hours is no longer available to the operating crews. The second shift can see what the first shift has done on the daily control chart and compare their work. The third shift can see what the first and second shifts have done and compare their work. The first shift, however, never did get to compare their record with the other two crews.

A Shift Summary Chart (Figure 2) was, therefore, designed to maintain at each press the total quality information of the job running until that job runs off.

In the Quality Control inspectors' schedule there is a half hour overlap at the end of each shift. During this period the two inspectors review the jobs running and the quality problems. It is the task of the inspector coming on duty to summarize, during this overlap period from the daily control chart at each press, the total severity points per carton for the shift, along with the number and types of defects found. This information is posted on the Shift Summary Chart. The Shift Summary Chart, therefore, indicates trends that exist during the job. It points out differences that may exist between crews or inspectors and it indicates which defects are most responsible for the quality record. Limits for the Shift Summary Chart are determined by $\overline{SP/c} \pm 3\sqrt{\frac{\overline{SP/c}}{kN}}$ where k is the number of cartons on the printed sheet and N is the number of sheets sampled per shift.

There are, of course, many repeat orders on a large number of items. The Job Summary Chart was designed (See figure 3) to help us appreciate from job order to job order

on a given item the status of control with time. When the job order runs off, the Shift Summary Chart is sent to the Quality Control office and the data from it summarized and entered on a Job Summary Chart.

The purpose of this chart is to help us determine our long range ability to maintain a quality level on a given carton. It helps us detect trends toward better or worse quality as repeat orders are processed. It also summarizes the type and severity of defects responsible for the quality level.

Job Summary Charts are reviewed by the superintendent before each new job order is run on a repeat item. The control limit formula for the Job Summary Chart is:

$$\overline{SP/c} \pm 3\sqrt{\frac{\overline{SP/c}}{kT}}$$

where k is the number of die positions on the sheet and T is the total number of tests.

To appreciate the effect of quality changes on a department basis, the Department Severity Points/Carton Chart was designed. The daily and monthly Printing Department record is based on the aggregate effect of all jobs running. Figure 4a is the chart of monthly averages as it appears in the monthly Quality Progress Report. Figure 4b is the chart of daily averages for the period covered by the monthly report.

Two sigma limits are used on the monthly chart and three sigma limits on the daily chart.

Figure 4a illustrates a shift in the process level which resulted from the deliberate introduction of an assignable cause. In May 1956 it was decided that protecting loads of board from changes in humidity might result in fewer defects due to wavy board. At that time the department SP/c record had been maintained for seventeen months. The process mean and standard deviation were calculated. Thus far the process seemed reasonably stable and two sigma rather than three sigma limits were used to give us a better chance (with greater risks) of detecting a shift in the process level.

June was out of control and also the best quality record since this chart was begun. July and August were also out of control on the low side and by the end of August we were pretty well convinced that a permanent shift in the process level had taken place. September's average made us wonder whether the beneficial effect was only to be gained during the hot summer months. October's average did little to dispel that doubt. However, November and December's averages indicate the beneficial effect of the assignable cause introduced is a permanent one and independent of seasonal fluctuations.

A potentially serious quality problem in the Printing Department is offset. Offset, in this case, means the transfer of ink from the surface of a printed sheet to the bottom of the sheet above it in a load of printed sheets. This occurs when the ink does not dry properly. When the loads are repiled several hours after printing, a sheet is selected a few inches from the bottom of each load and inspected for this offset condition. The pressmen's identification is on each sheet. The offset condition can, therefore, be charged back to the right shift and pressman although the repiling may take place on subsequent shifts. A quality record is maintained by individual pressmen and by crew as well as department.

After keeping the offset record by crew for a few months, it was noted that although the three crews were usually quite close together, crew A was consistently showing the best percent defective offset. Was crew A significantly better? We thought so. It was felt that if the difference between crew A and the other crews was significantly different, even if the absolute difference seemed relatively small, that the possible existence of a favorable assignable cause could be investigated.

Our null hypothesis was that all crews are the same and that apparent differences are not real differences. Let a equal the number of ways in which crew A can have the worst record. Let b equal the number of ways in which it does not have the worst record. a and b are mutually exclusive and equally likely to occur. There are three crews. Therefore, a = 1 and b = 2.

The probability of a is $a/a+b = 1/3$

The probability of b is $b/a+b = 2/3$

At the end of eight months it was noted that crew A had had the worst record only one month, crew B three months and crew C four months.

We wished to determine then the probability of crew A having the worst record for one out of eight months when the probability of the worst record for any one month is 1/3.

$$P\left(\frac{c}{N}\right) = \frac{N!}{c!(N-c)!}\, a^c b^{N-c}$$

$$P\left(\frac{1}{8}\right) = \frac{8!}{1!\,7!}\,(0.33)^1(.67)^7$$

$$= (8)(0.33)(.06)$$

$$= .15$$

This was not a sufficiently remote probability to be significant.

However, at the end of twelve months, there was still but one month when crew A had the worst record.

$$P\left(\frac{1}{12}\right) = \frac{12!}{1!\ 11!}\ (0.33)^1(.67)^{11}$$

$$= (12)(0.33)(.0122)$$

$$= .048$$

This seemed fairly significant. Only 4.8% of the time could such a record have occurred due to chance cause. Investigation indicated that operating techniques were consistent between shifts. Crews rotate once a week between shifts so there is no reason for crew preference. The distribution of good pressmen seemed balanced between crews.

It was then asked, "Is the foreman of crew A a better supervisor? If so, what does he do differently to control offset that might be incorporated in the supervisory practices of the other two foremen?" Attention was, therefore, directed to the foremen. Another twelve months passed. At the end of twenty four months, crew A had had the worst record for three of those months.

$$P\left(\frac{3}{24}\right) = \frac{24!}{3!\ 21!}\ (0.33)^3(.67)^{21}$$

$$= (2024)\ (.0359)\ (.000223)$$

$$= .0162$$

1.62% of the time such a record could occur due to chance cause. This seemed to be sufficient proof of a significant difference. There was also evidence of better supervisory practices by the foreman of crew A.

Cutting

Quality Control procedures in the Cutting Department are essentially the same as in Printing. Only the defects have been changed. There are, of course, a larger proportion of functional defects, many of which can only be evaluated visually. One thing different is the approach to stripping defects. Stripping is the act of removing scrap from the sheets of cut cartons. The nature of this act is such that most stripping defects are critical defects. Only critical defects are, therefore, tabulated. Stripping defects are limited to torn cartons, hammer marks, wrinkles or unsanded nicks.

The sample size is based on a "lift" of cartons. A lift is the number of sheets pulled at one time from the cutting press by the strippers to be hand stripped. To determine the number of cartons in a lift, a multiple pocket scale is placed along side the lift. To determine which is

the applicable scale for the job order on a press, 50 cartons are chosen at random at the beginning of the run. The total caliper of these 50 cartons are compared with the multiple scale. The number 50 on the particular scale which comes closest to the total caliper of the 50 cartons is the scale used throughout the rest of the run. A lift contains from 100 to 300 cartons. The nature of stripping defects also makes it possible to detect these defects by fanning rapidly through the lift. The control chart used is the p chart (figure 5). While lift sizes may vary, they do not vary enough to prohibit using the conventional p chart.

Gluing

Quality Control in the glue room is a combination of in-process Quality Control for the gluing operation and an out-going quality check on printing, cutting and gluing operations.

At the present time there are two different types of sampling procedures in the glue room quality check. One is a sampling with time, the other by units of production. The first is a continuous quality audit which requires subsequent changes in the stringency of production screening to maintain the A.Q.L. The second considers two or three consecutive containers of cartons as a lot and employes a lot by lot acceptance sampling plan to maintain the A.Q.L. Rejected lots are 100% sorted.

Two types of control charts are used in the glue room. One is the combination Severity Points/Carton chart (Figure 6) for printing, cutting and gluing defects. The other is a cumulative fraction defective control chart (Figure 7) which plots a defective quality index based on defectives rather than defects or severity points.

On the combination Severity Points/Carton chart a double graph line is plotted on the outgoing quality recheck of printing and cutting defects but without control limits. On the gluing defects chart graph grid there are limits.

The Cumulative Defective Chart was designed to supplement or substitute for the conventional fraction defective or percent defective control chart. It makes use of cumulative sample size and the result is increasing reliability of prediction of lot fraction defective and relationship between this fraction defective and statistical control limits.

Only computations involving addition are required by the user of this chart. The point last plotted reveals the number of defectives thus far, the fraction defective of the cumulative sample, and the relationship of the lot to control limits. The control limits come ever closer to the A.Q.L. as the cumulative sample size increases. This is due

to the fact that as a sample size increases, its average tends to come closer to the true population average.

Points inside control limits vary from random causes if the subgroup samples are random. A point outside of control limits presses the conclusion that the conditions which causes the process to deliver at the A.Q.L. no longer exist. It is required that either corrective action be taken to re-institute the former set of conditions or accept a new A.Q.L. with different limits.

The chart is adaptable to variations in lot size and sample size, and can be readily applied by personnel of limited background.

Glue Lap-Score Caliper Control

An application of the use of an $\bar{X}$ and R Chart has been the control of caliper relationship between glue lap and parallel 180° scores of glued cartons.

The customers' specifications call for the caliper of the far edge of the carton to average the same as the glue lap with ± .001 inches tolerance limits for sample averages. The near edge of the carton was to caliper .006 inches less than the glue lap with the same tolerance limits. Sample size was twenty cartons.

For control chart purposes, a sample of n = 5 cartons was chosen. The relationship between standard deviations and sample sizes is:

$$\frac{\sigma_{n_2}}{\sigma_{n_1}} = \frac{\sqrt{n_1}}{\sqrt{n_2}} \quad \text{where } n_1 = 20 \text{ and } n_2 = 5.$$

This equation states that the ratio of the standard deviations varies inversely as the ratio of the square roots of the sample sizes:

$$\frac{\sigma_{n_2}}{\sigma_{n_1}} = \frac{\sqrt{20}}{\sqrt{5}} = \frac{4.470}{2.235} = 2$$

$$\sigma_{n_2} = 2\sigma_{n_1}$$

This tells us that if the averages of five cartons stay within ± .002 inches that the averages of twenty cartons will stay within the specified limits of ± .001 inches. Could we meet those specifications? First, it was necessary to determine process capability.

The first $\bar{X}$ and R Chart (Figure 8) indicated that glue machine adjustments could be made to bring the average caliper where specified and that the specified limits coincided

quite closely with process capability. In fact, spec limits are a little looser than statistical limits based on normal process variability.

Defect Standards

At a central location in the carton plant there is maintained a file of standard defects. These standard defects represent categories of items which have been classified according to carton plant department similarity of quality requirements, style and copy. The ideal situation would be a three part specification book for each item, or one book for each carton plant department due to the size of the book. This would mean, however, hundreds of defect books. The items were, therefore, categorized.

There are five to ten categories in each carton plant department. Representative defects selected from each category are a guide to the severity evaluation of the Quality Control Department. Disputes between production and Quality Control can, therefore, be referred to these files. The defective condition in dispute may not have a file defect from that exact area on the carton or perhaps even from that carton but the range of standard defects in that category usually leaves little doubt as to the severity of the defective condition in dispute.

While standard defects tie down specific defective conditions, visual evaluation of quality is of necessity a mental concept. This concept is best expressed in the definition of the three categories of severity. Inspector experience is, therefore, important. This concept must be learned, become ingrained, because if every defective condition found had to be referred to the standard defects for comparison then very little inspection would be accomplished. Far more inspection and, therefore, control is attained by using the defect standards principally to maintain the stability of the inspector mental evaluation level. It is better to tolerate variation around this level as a result of subjective evaluation than to circumscribe the sample size by insisting on exact comparisons with the standards.

The consistency of subjective judgement has been developed so well by our experienced inspectors that a new job starting up for the first time can be evaluated immediately and accurately in terms of the category of items which it fits. Later, decision by both production and quality control management as to the severity of standard defects of the new item almost always finds complete agreement with the inspectors' judgement. Our long range goal is always to replace judgement with objective test methods of measurement. However, should the time ever come when every quality characteristic can be measured objectively, I doubt that the nature of our product would permit us the luxury of a large number of exacting physical tests on every carton of every sample.

Periodic Quality Reports to Management

1. Daily

Daily reporting to carton plant management of the progress of quality is implemented by:

a. daily control charts
b. representative samples and defects
c. daily and cumulative summary by job (Figure 9)

The Daily and Cumulative Summary sums up for every 24 hours on each job that is running the total number of major and critical defects by type of defect. On some jobs control by individual defects is more important than the aggregate effect of all defects. This summary permits a daily revaluation by individual defect and total defects. Circling of listed values on the report indicate to production management when selected AQL or target maximum values have been exceeded.

2. Monthly

The monthly Progress Report summarizes all inspection data month by month for a period of two years or more so that long range trends can be appreciated. It also details the daily picture during the month covered by the report. Figures 4a and 4b are one type of chart which appear in the report. As much as possible, all data is reduced to graphs.

Monthly averages are maintained on a master graph. Only one point is plotted at the end of the month for each control characteristic. The master graph is then photographed and transferred to a printing plate as part of the multilith office reproduction process. Because of the use of master graphs only the most recent issue of the monthly Progress Report need be filed. Each issue summarizes all past data.

One purpose of the Progress Report is in analyzing and reporting board mill data to aid in determining the accuracy of board specifications. For example, the weight-caliper relationship of paperboard grades is extremely important. Failure to meet the specified basis weight for a given caliper is usually due to one of these reasons:

1. inaccurate specifications
2. short term inability of operating crews to meet specifications
3. deliberate departure from specifications

The Progress Report grade run analysis helps establish accurate and realistic specifications.

Special Projects

1. Analysis of Variance

In a study in the Cutting Department, cutting knives of different material were used under various conditions to determine which type of metal gave the best wear. In order to compare these metals it had to be determined that each test was sufficiently homogeneous as to indicate that a good normal makeready had been performed on the press and, therefore, valid comparisons could be made between runs.

Three rows of measurements were made around and across the cylinder. Analysis of variance (Figure 10) on one press indicated that conditions from bearer to bearer across the cylinder were uniform and that conditions around the cylinder were not. This meant that before around-the-cylinder variation could be checked on 65 chrome metal on printed board against another metal on printed board that a new and better makeready would have to be done. That part of the experiment would then have to be repeated. Apparent differences between metals were then tested for significance with the t test.

2. Linear Regression

A grade of board is made at the mill with an asphalt center as a moisture barrier to protect the customer's product. To control the amount of asphalt the water vapor transmission rate must be determined. As a laboratory test this is slow and relatively expensive. A possible method of measuring this asphalt layer is an air permeability test which can be run in the mill quickly and conveniently. A series of measurements indicated that there was a relationship between air permeability and WVTR (Water Vapor Transmission Rate). How good was this relationship? Was it linear? Could a predictive equation be derived which could be used to effectively control the thickness of the asphalt layer within the WVTR specification limit?

A series of 140 readings were made with air permeability as the independent variable and WVTR as the dependent variable. Of these 140 readings 136 formed a linear, relatively tight scatter diagram. There were four extreme values which were discarded as a result of a randomness test. The 136 points were homoscedastic and fell within $\pm 3\sigma$. All but five fell within $\pm 2\sigma$ and those five were just outside $\pm 2\sigma$.

The maximum specification limit for WVTR was 5.0 (see figure 11). It was felt that 2.5% of values exceeding this maximum could be tolerated -- $\pm 2\sigma$ limits were, therefore, used. The intersection of the WVTR specification limit with the upper 2σ control limit indicates the upper limit of the air permeability value.

The relationship between air permeability and WVTR was

found to be good enough to replace the laboratory tests. The rapid air permeability tests in the mill reduced, therefore, the expense of testing and resulted in more effective control in the mill during manufacturing.

3. Key Sorter Punch Card File

Board mill data from standard board grades which comprise most of our tonage is summarized by order and entered on Key Sorter Punch Cards (see figure 12). The storing of data in this manner enables us to analyze board characteristics quickly and conveniently on any individual or combined basis of listed quality characteristics.

A typical example of a request for information might be: "How much of the time on all orders of Bleached Manila that ran from June to December on machines #4 and #5 were the run basis weight averages above and below specifications?" or "During 1955 and 1956 what was the caliper-stiffness relationship on machine #3 on all runs of White Patent Coated Solid Kraft Back?"

On the first request the needle first rejects cards from all the machines but #4 and #5. It then rejects all the grades except Bleached Manila. Next it eliminates all the dates except from June to December. The basis weight information on the remaining cards, therefore, can be tabulated and the number which fell above and below the specified basis weight determined.

On the next request the needle first isolates machine #3 data cards. It next selects White Patent Coated Solid Kraft Back grades. It next rejects all dates but 1955 and 1956. A scatter diagram and subsequent regression analysis can then be made between caliper and stiffness on the grade selected.

Industrial Relations Problems of Quality Control

We feel that the greatest problems still facing Quality Control today in our company are the personal rather than the mechanical aspects of the quality program. It is felt that the control charts, statistical and other analytical techniques and reporting procedures are effective and the right approach for our particular problems. However, there is and may always be an underlying antagonism from a few people who resent the objective publication of quality information. One foreman told me, "During the war we had to accept people for pressmen who weren't really pressmen. Then Quality Control came along after the war and forced them to be the pressmen they said they were and for which they are getting paid."

To the man on the press we are sometimes a system which has no heart, which does not sympathize with his temporary

problems which he feels are at times beyond his control. We are not fair, he feels, in maintaining a spotlight on his quality record. Both he and his foreman sometimes feel that we are the cause of strained relations between them.

Perhaps so. The Quality Control system calls for defect accountability. The Quality Control system forces a decision as to whether a defective condition should be corrected or allowed to continue -- in fact, a continuing pattern of decisions as to whether the customer will or will not accept a defective condition and if so, how much.

Some pressmen wonder why a category of defect severity which is defined as one "that which will always be accepted by the customer" needs to be marked up at all.

Superintendents sometimes feel that action taken by operating people is designed less to meet the optimum cost-quality picture than to make the control chart record look good. One superintendent has stated that pressmen may be relying too much on Quality Control to find defective conditions and this results in the pressman not checking accurately enough or frequently enough his own work.

There are always, of course, old timers who feel it is popular to dislike a staff function such as Quality Control. Interestingly enough these people are usually the most effective users of the Quality Control system. Their attitude is not unlike that of a twelve year old boy who was in love with a girl but would have died before admitting it.

These Industrial Relations problems do exist. However, most operating personnel want and like Quality Control. Quality Control is different things to different people. How different people feel can be best expressed in their own words.

Pressman - "I depend on the inspectors a lot especially when I have a makeready on one press. They give us a lot of information. These are the best girls now we have ever had. It is a very rare occurance when they miss anything."

Foreman - "Things are going pretty well between us (production and Quality Control). I don't think they can do much more unless they get a hell of a lot more training than they are ever going to be able to get. They do find a lot of stuff that we don't. Of course, we might find it sooner or later but it is better to have it found sooner so that we don't make anymore bad stuff."

Pressman - "I don't mind admitting that there has been times when they have saved my neck. It has also saved the necks of others whether they admit it or not. There have been many times when they have found things that I hadn't noticed. To me, Quality Control may be the difference in

letting 200 or 300 or 2,000 or 3,000 bad cartons run. They've got their job to do and I have got mine. Some of the guys in here who said they don't like women coming in and telling them what to do and they don't like Quality Control are the first ones to yell when the Quality Control girl misses anything."

Pressman - "Especially when I am busy on one press and don't have time to take many tests on the other one I depend on them a lot. I have told them to bring me information of serious defects just as soon as they can and they do it. That is a big help. Some of the men think they mark too strict but it is their job to find it and mark it up. They are a big help to me."

Pressman - "You have a G-- D--- good bunch of girls here and they are a big help. Why there has been many times that if they hadn't brought a sheet back and showed me where something was I might have run a 1,000 sheets or more before it was found. My eyes aren't what they use to be and her bringing the sheets back with the defects circled is a big help. They really give those sheets a going over."

Pressman - "Your job is to find and get the information and get it to us. Thereafter it is up to us to do something about it. I don't know what else you can do. Sometimes the guys feel the girls think they have to find something on every sheet but of course, it is pretty hard to not have something on every sheet. You have got sheet to sheet variation, thickness, smoothness and so forth and it would be pretty hard not to have something show up as a defect."

Foreman - "The girls do a lot of good. We need them out here and couldn't do without them. Sometimes I get pretty busy. The other day, for instance, I had two or three okays in a row and if it wasn't for the fact that the girls make those periodic checks a lot of things would get by.

Frequently, all I have to do is walk down the aisle and take a look at the control chart and if I see something wrong I can immediately go and check out the source. There was one job where unsanded nicks and torn cartons were found by the girls. An inspector had found three okays in a row. All of a sudden on her fourth test the graph line took a big jump up and when I saw that I went immediately to the source of the trouble and found that the strippers had not been following instructions. I shut the press down right there and then and had an education session with them. Now if that graph line hadn't been there to show me I would have probably walked right on by and not gone over to check."

Pressman - "Take today, for example. We were breaking in a new feeder and it was breaking us in too. There were a lot of things wrong on the sheet. Trying to get the bugs

out of this feeder we don't have time to check the sheet very carefully. It is a big help to have them come and look at the sheet and know that everything has been marked down in detail. You have got a good bunch of girls here."

Foreman - "Before there was Quality Control the foreman might have seven or eight presses running. Maybe three okays one right after another. It might take him two hours or so to get back to the press running in the corner. Now the foreman knows the presses are being checked regularly by Quality Control inspectors and that makes it a lot easier for the foremen."

Pressman - "They (the inspectors) are an extra pair of eyes."

Summary

In this discussion there has been an attempt to illustrate the use of Quality Control in the paper industry as it is practiced at The Gardner Board and Carton Company. Management's philosophy of Quality Control and some of the control techniques have been highlighted. Some of the Industrial Relations problems which must be overcome in selling Quality Control to people so that they will use it and profit by it have been detailed. Our program is in a constant state of flux. Control charts are being refined, new techniques investigated and adopted, and reporting procedures improved.

Most distributions of data with which Quality Control is concerned follow mathematical laws. This makes possible the development of scientific procedures for the control of quality.

Quality Control is concerned with people, therefore, an understanding of the fundamental nature of people is necessary if a Quality Control system is to be successful.

In our company we have tried to understand the basic nature of data. We have also tried to understand the basic nature of people. If our Quality Control program has been successful it is because of our attempt to understand and integrate these factors -- plus the fact that quality is a company policy.

(Figure D)

PRINTING DEPARTMENT

ORDER NO M-4226 CUSTOMER Dow Brand Co. - Soap DEFECTIVE QUALITY LEVEL REPORT DATE 2-21-57 PRESS NO 16

REMARKS

1ST SHIFT
#1 Off. Off. & ... F.R. ...

2ND SHIFT

3RD SHIFT

COUNTER READING 4:00 8:00 10:00 12:00 2:00 4:00 6:00 8:00 10:00 10:00 12:00 2:00 4:00 6:00

RUN 1
N 12
COLORS

SEVERITY POINTS/CARTON 4.0 3.0 2.0 1.0 0

Average

DEFECTS

	4:00	8:00	10:00	12:00	2:00	4:00	6:00	8:00	8:00	10:00	10:00	12:00	2:00	4:00	6:00
1 OFF REGISTER	5 7									10			10	10	
2 WASHED PLATE														(4)	
3 INK FILL-UP	(2) (12)	4	4 8	4 9	1 4			1 2 5 9	8 9			3 7			(7)
4 DIRTY FORMS AND SPOTS	5 9 (9) (12)							(11)	1 3	(3)	8				
5 SMEARS AND STREAKS	(3)	(7)	9	1											
6 TRAP	9											9	9	9 10	9 10
7 B.U.	9 8				6 7				9 11		7			9 7	6 4
8 POWDER		(6)	8		6										
9 SGM.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
TOTAL SEVERITY POINTS	19	18	3	2	3	3	2	4	2	5	2	4	3	0	2	1	0	2	8	2	4	1	0

DEFECTS OKAYED BY FOREMEN — NOT CHARGEABLE TO PRESSMEN

1 OFF Reg.	5 7	5 7	5 7	5 7	5 7	5 7	5 7
2							
3							
4							
TOTAL SEVERITY POINTS	2	2	2	2	2	2	2

WHITE - PAGE HADDIX - COGAN Cooper - Monitor

(Figure 2)

PRINTING Quality Control Shift Summary Chart

Press No. 16 Run 1 N 12 Order No. M-4238 Customer Andras

Carton Description Napkins

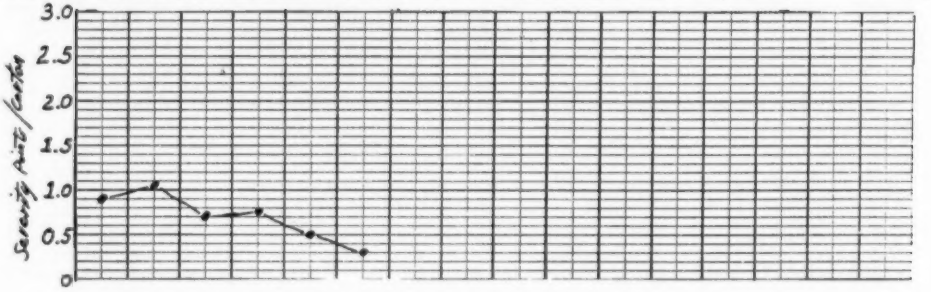
| | | |

(Figure 3)

Customer JOB SUMMARY CHART for Printing Department

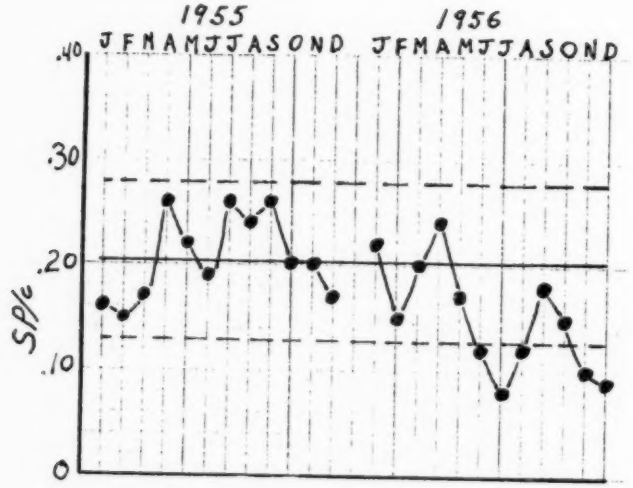
Customer and Job Description Andros Co. NAPKINS

Order No.	M-2133	M-2189	M-4418	M-6922	M-7702	M-8693										
Date In	11-3-55	2-12-56	5-28-56	8-14-56	12-3-56	2-12-57										
Date out	11-11-55	2-22-56	6-6-56	8-23-56	12-14-56	2-27-57										
No. Cartons Inspected	1770	1630	1510	960	1280	1420										

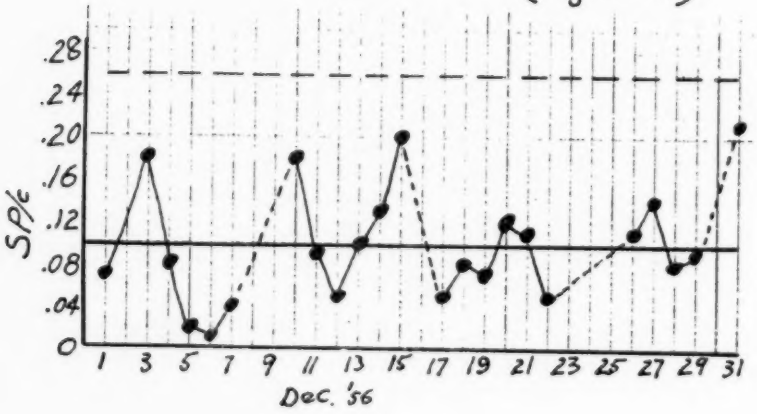


DEFECTS % OF	#3	#5	#3	#6	#3	#5	#3	#6	#3	#5	#3	#5	#3	#5	#3	#6	#3	#6	#3	#6	#3	#6	#3	#6	#3	#6	#3	#6	#3	#6
Color Reg.	6.0	.6	3.3	2.1	3.8		2.7	.6	1.2		.6	.1																
Mesh	.4	.02				.5		.3			.3																	
Dirty Form			.7	.04		.5	1.1				.1																	
Fill Up	2.8	.6	1.5		1.3		.4	.3	2.3	.2	.08																	
Dead Metal				1.1				.05	.06																			
Oil Spots		.03		.2																								
Light Imp.	1.7				.4	.8	.6		.2			.1																	
Sticky & Smy.		.01			.06			.9		.03																		
S.G. Marks																												
Def. Board			2.2		1.4						.06																	

Printing SP/c (Figure 4a)

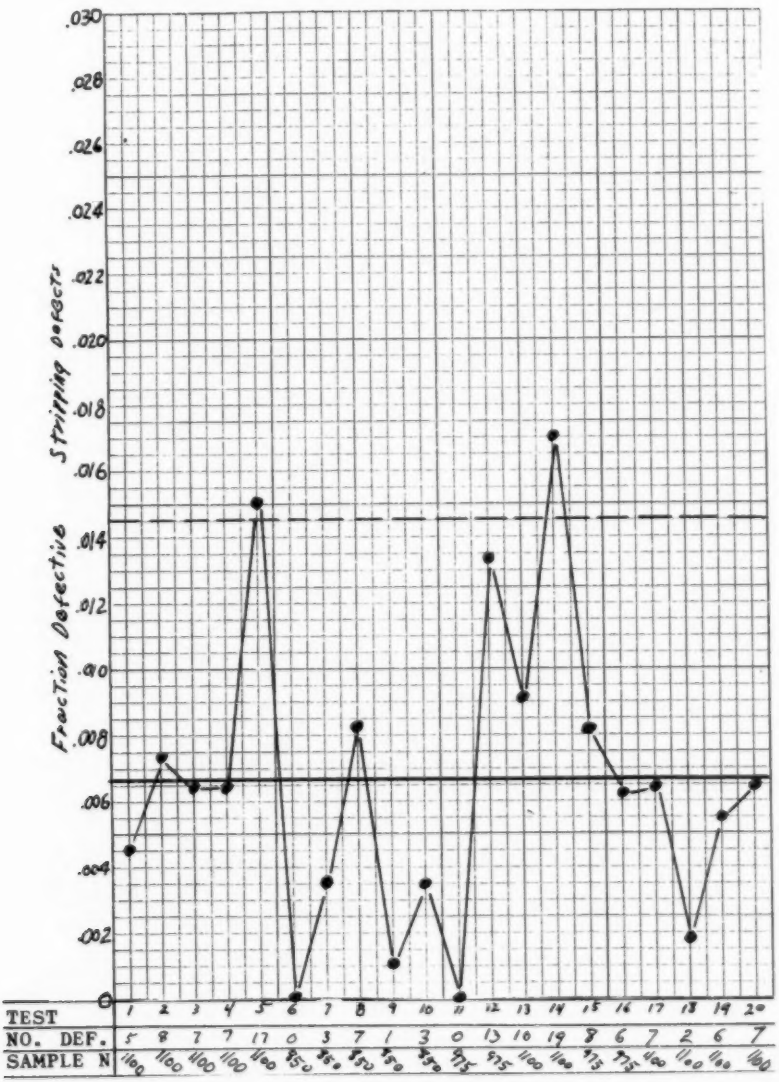


(Figure 4b)

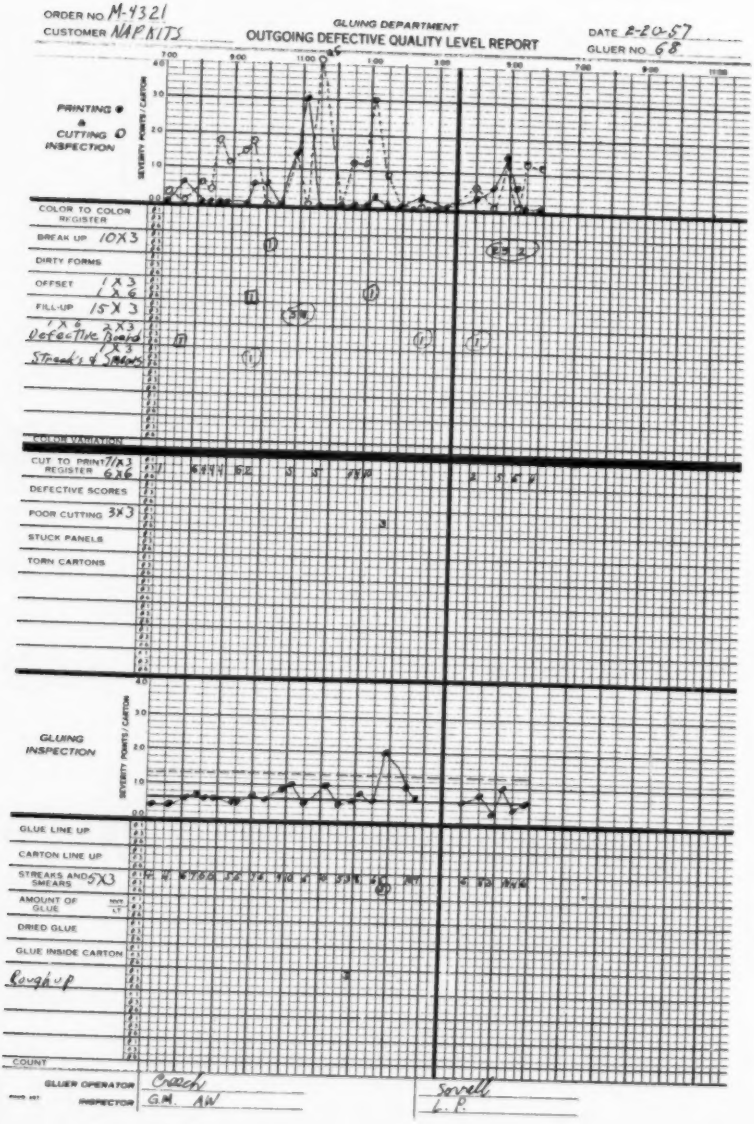


(Figure 5)

DATE: 5 Mar '57 ITEM: Nat'l Tob. Co.'s REFRESH



(Figure 6)


ORDER NO. M-4321
CUSTOMER NAPKITS
GLUING DEPARTMENT
OUTGOING DEFECTIVE QUALITY LEVEL REPORT
DATE 2-20-57
GLUER NO. 68
PRINTING ●
& CUTTING ○
INSPECTION
SEVERITY POINTS / CARTON
COLOR TO COLOR REGISTER
BREAK UP 10X3
DIRTY FORMS
OFFSET 1X3 2X6
FILL-UP 15X3
1X3 2X3 Defective Board 1X3 Streak's & Smears
COLOR VARIATION
CUT TO PRINT REGISTER 7/X3 6X6
DEFECTIVE SCORES
POOR CUTTING 3X3
STUCK PANELS
TORN CARTONS
GLUING INSPECTION
SEVERITY POINTS / CARTON
GLUE LINE UP
CARTON LINE UP
STREAKS AND SMEARS 5X3
AMOUNT OF GLUE
DRIED GLUE
GLUE INSIDE CARTON
Roughup
COUNT
GLUER OPERATOR
INSPECTOR
Veach
G.M. AW
Sorrell
L.P.

(Figure 7)

Student Number: M-2131 Customer: KODAK Date: 16-1-57

Cumulative Fraction Defective Control Chart

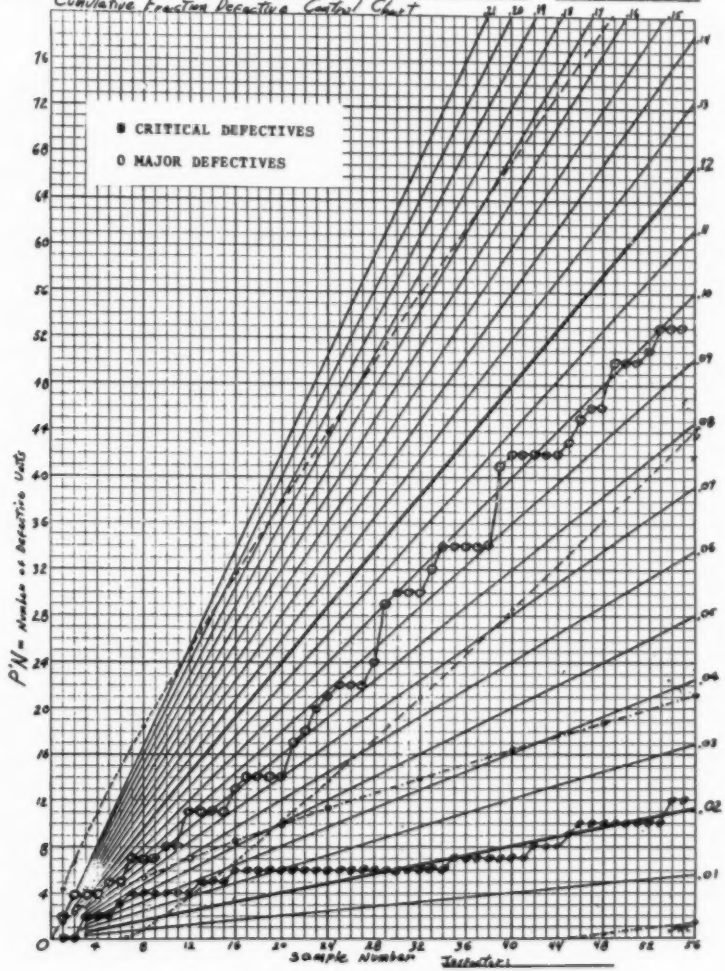
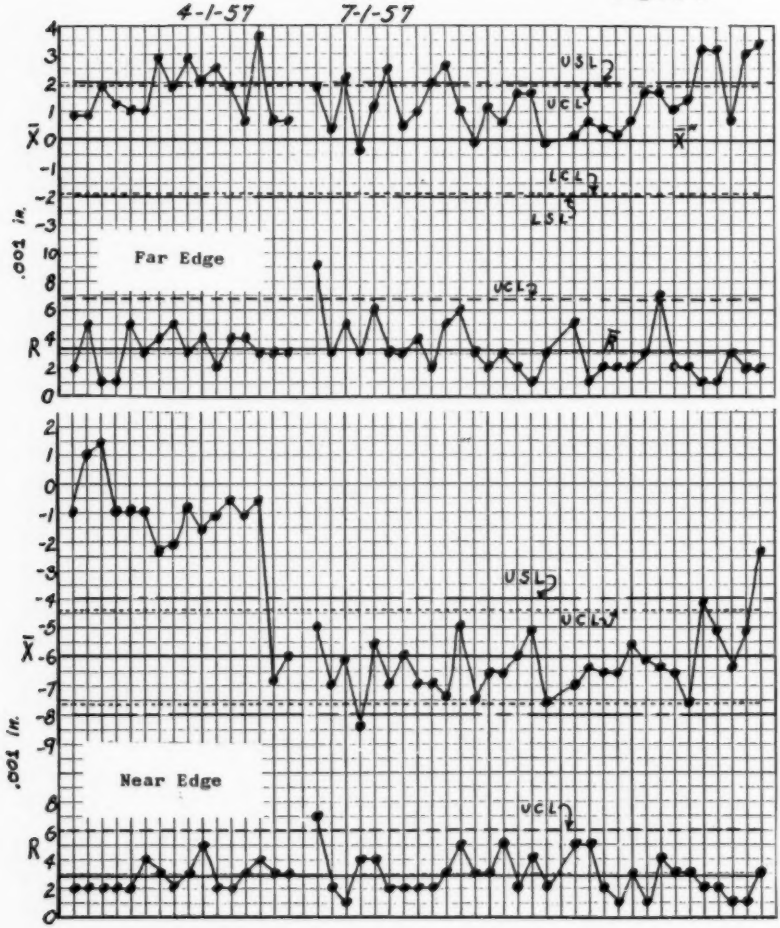


Figure 8



(Figure 9)

MIDDLETOWN QC DAILY AND CUMULATIVE QUALITY RECORD

DATE: 15 Mar '57

Defects	SOAPY SUDS				NAP KITS				KOPAK			
Printing	Today N=160		Cum. N=510		Today N=48		Cum. N=48		Today N=380		Cum. N=920	
	#3	#6	#3	#6	#3	#6	#3	#6	#3	#6	#3	#6
Register	6	2	10	10	(3)		(3)		16	3	38	5
Dirty Forms			4						4	1	4	1
Fill Up									15	(8)	(72)	17
Ink Break Up	(15)	(6)	(28)	9								
Mesh					(6)		(6)			1		2
Dead Metal	1		1							6		8
Streaks & Smears			5									
Trap					2	1	2	1				
Total Printing	22	8	48	19	11	1	11	1	35	19	154	33
Average (Defects/carton)	.14	.05	.09	.04	.23	.02	.23	.02	.09	.05	.17	.04
Offset	9	2	24	10					18	2	33	5
Percent Offset	2.0	0.4	2.0	0.8					2.3	.25	1.1	.16

Defects								
Printing	Today N=		Cum. N=		Today N=		Cum. N=		Today N=		Cum. N=	
	#3	#6	#3	#6	#3	#6	#3	#6	#3	#6	#3	#6
Register												
Dirty Forms												
Fill Up												
Ink Break Up												
Mesh												
Dead Metal												
Streaks & Smears												
Trap												
Total Printing												
Average (Defects/carton)												
Offset												
Total Offset												
Percent												

(Figure 10)

ANALYSIS OF VARIANCE

65 Chrome on Ink

Across the cylinder

Row 1	Row 2	Row 3
1	2	1
1	2	2
2	2	2
.	.	.
.	.	.
.	.	.
4	3	4
2	1	2
2	1	4
ΣX = 55	46	47 : 148
ΣX^2 =157	100	123 : 380

N = 24

Total S.S. = $380 - \frac{148^2}{72} = 75.7778$

Row S.S. = $306.2500 - \frac{148^2}{72} = 2.0278$

Source of Variation	Sums of Squares	Degrees Freedom	Mean Square	F.	F.05
Total	75.7778	71			
Row	2.0278	2	1.0139	.9486	19.47
Error	73.7500	69	1.0688		

Around the cylinder

Row 1	Row 2	Row 3
2	2	3
1	3	1
2	3	3
.	.	.
.	.	.
.	.	.
4	3	5
2	4	4
2	4	4
ΣX = 54	75	88 : 217
ΣX^2 =134	275	370 : 779

N = 24

Total S.S. = $779 - \frac{217^2}{72} = 124.9861$

Row S.S. = $678.5417 - \frac{217^2}{72} = 24.5278$

Source of Variation	Sums of Squares	Degrees Freedom	Mean Square	F.	F.05
Total	124.9861	71			
Row	24.5278	2	12.2639	8.42	3.15
Error	100.4583	69	1.4559		

(Figure 11)

X	Y
40	3.4
22	2.1
..	...
..	...
..	...
83	4.6
59	2.5

N = 136

ΣX =	$(\Sigma X)(\Sigma Y)$ =	$(\Sigma Y)^2/N$ =
ΣY =	$(\Sigma X)(\Sigma Y)/N$ =	$\Sigma xy = \Sigma XY - (\Sigma X)(\Sigma Y)/N$
ΣXY =	$(\Sigma X)^2$ =	$\Sigma x^2 = \Sigma X^2 - (\Sigma X)^2/N$
$\bar{X}$ =	$(\Sigma X)^2/N$ =	$\Sigma y^2 = \Sigma Y^2 - (\Sigma Y)^2/N$
$\bar{Y}$ =	$(\Sigma Y)^2$ =	

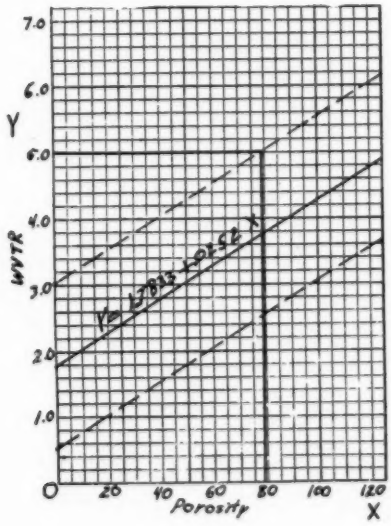
$$b = \frac{\Sigma xy}{\Sigma x^2}$$

Estimating equation = $Y_e = (\bar{Y} - b\bar{X}) + bX$

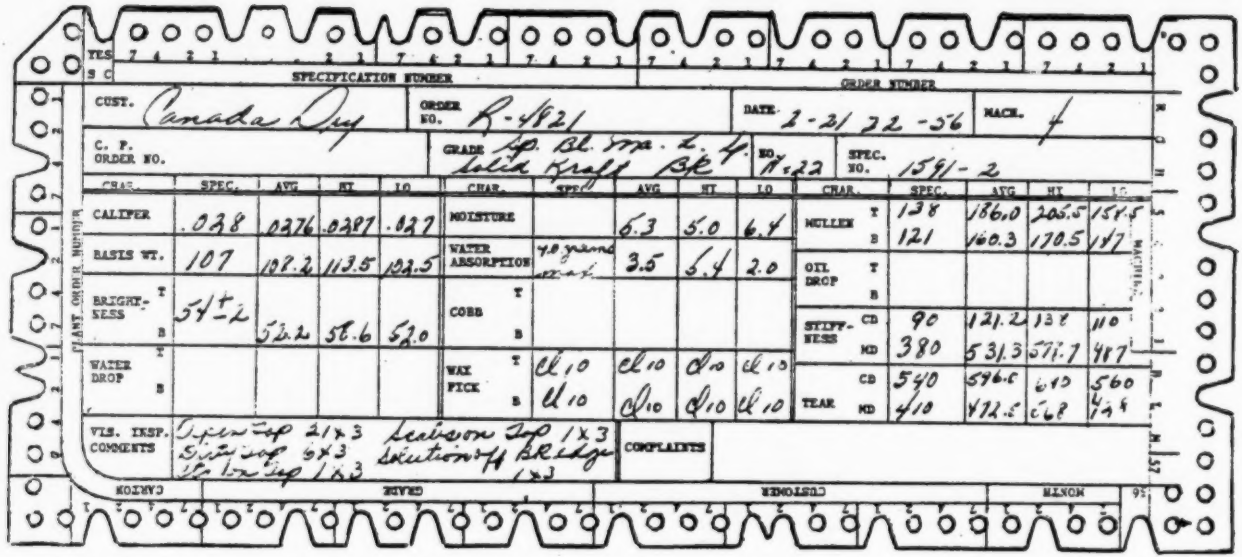
$$Y_e = 1.7833 + .0252X$$

An estimate of the variance about the regression line is given by

$$\sigma^2 = \frac{\left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}\right] - b^2\left[\Sigma X^2 - \frac{(\Sigma X)^2}{N}\right]}{N - 2}$$



(Figure 12)


SPECIFICATION NUMBER
ORDER NUMBER
CUST. Canada Dry
ORDER NO. R-4921
DATE 2-21 22-56
MACH. 4
C. P. ORDER NO.
GRADE Lp. Bl. Mx. 2.4 Solid Kraft BK
NO. R-22
SPEC. NO. 1591-2
PLANT ORDER NUMBER
CARTON
CRATE
CUSTOMER
MONTH
CHAR. | SPEC. | AVG | HI | LO
CALIPER | .028 | .0276 | .0287 | .027
BASIS WT. | 107 | 107.2 | 113.5 | 102.5
BRIGHTNESS T | 54±2 | | | |
BRIGHTNESS B | | 52.2 | 56.6 | 52.0
WATER DROP T | | | | |
WATER DROP B | | | | |
MOISTURE | | 6.3 | 5.0 | 6.4
WATER ABSORPTION | 40 grams max. | 3.5 | 6.4 | 2.0
COBB T | | | | |
COBB B | | | | |
WAX PICK T | | Cl 10 | Cl 10 | Cl 10 | Cl 10
WAX PICK B | | Cl 10 | Cl 10 | Cl 10 | Cl 10
MULLEN T | 138 | 186.0 | 205.5 | 154.5
MULLEN B | 121 | 160.3 | 170.5 | 147
OIL DROP T | | | | |
OIL DROP B | | | | |
STIFFNESS CD | 90 | 121.2 | 138 | 110
STIFFNESS MD | 380 | 531.3 | 571.7 | 487
TEAR CD | 540 | 576.5 | 610 | 560
TEAR MD | 410 | 472.5 | 568 | 414
VIS. INSP. COMMENTS: Open Top 21x3 Division Top 1x3 Partition 9x3 Bridge 1x3 Bottom Top 1x3
COMPLAINTS

FACTORS AFFECTING THE ACCURACY AND RELIABILITY OF SENSORY TESTS

David R. Peryam
Quartermaster Food and Container Institute

Prediction of population behavior toward foods and controlling food manufacture in light of such behavior are getting to be big business. The manifold activities that have to do with flavor, odor and acceptance play a vital role for many companies, particularly in the development of products and the planning of future operations. Traditionally this important area has always been the domain of the expert, but, as competition for the consumers' favor has become keener, interest in the scientific approach has grown. The idea appears ever more reasonable and economically sound that we can predict what consumers en masse are likely to accept, and that through sensory testing we can manipulate the flavor factors which are crucial to acceptance.

A great deal has already been achieved, but the impression is current, even among those who have successful programs, that this kind of testing is not on as sound a basis as it should be. If the precision of physical measurement is taken as a model, measurement in sensory testing may seem inaccurate and unreliable. This attitude may or may not be justified. There is considerable confusion in this area. People with many different scientific and experience backgrounds are working in it, are trying to find out many different things, and are using many different approaches. The field is not systematized and integrated. Thus, when inaccuracy and variability are noted, their causes and possible means of control are not immediately clear.

My purpose is general inquiry into the factors involved, but using review and discussion rather than exposition. The objective is to provide orientation, rather than to impart specific information. The latter will not be consciously avoided. However, I will not assume responsibility for providing solutions to the problems raised, but will simply try to get them into perspective. Throughout we will be concerned with measurement--measurement both of human behavior and by means of human behavior. Basically, then, we are in the field of psychology, but we shall try to maintain the applied viewpoint of one who is more interested in products than in the human behavior as such.

What kinds of deficiencies in sensory test results are alleged? Without being exhaustive we can mention the following: failure to discriminate differences in quality between products, disagreement among individuals or among groups in values assigned the same product, or changes in the values assigned by the same person or group of persons upon different occasions. Finally, we should note that sensory testing is often blamed simply because of failure of the test results to confirm the experimenter's expectations as based upon some validity criterion such as a known physical variable. Some of these reasons are certainly valid while others, e.g., the last, probably are not.

In this area is it reasonable to strive for the same order of precision that we expect in physical measurement? Both the possibility and the desirability of achieving this end are questionable.

The judged reliability of sensory tests may depend upon one's point of view. If our objective is to reproduce the same values as closely as

possible from time to time and place to place, we may conclude that they are unreliable. In this limited context the discrepancies are considered as error; however, we may otherwise consider the observed differences as indicating real changes in test conditions or test populations. Thus, they are important in their own right and give us essential information. The trouble is that usually we don't know what these conditions are; therefore, we don't know how to interpret the variability. Many of the measures we use must inevitably remain variable, hence subject to the suspicion of unreliability, because we use them to obtain information about phenomena which are innately variable. But the amount of this innate variability itself varies according to test objective and test form, so let's look at both.

In discussing types of tests a rather coarse grouping will be adequate for our purposes. Four types have been defined and are presented in Figure 1. You will note that one type is outside the scope of this paper, being labeled as "non-sensory". However, it is included because it will provide some useful comparisons. The distinction between sensory and non-sensory is easy; the term "sensory" is reserved for those types of tests which involve actual samples of food. They are contrasted with the survey type, where the subject responds to the idea of a food as represented by its name on a questionnaire. Typically, sensory tests belong in the laboratory, and surveys out in the field. The survey objectives are broader. They seek to measure learned attitudes and opinions. From them we may infer the behavior potential for types of products, i.e., how "popular" they are, but rarely can surveys provide specific information. You might say their purpose is to get at the characteristics of populations. In consequence, their value is markedly reduced unless they are conducted on a fairly broad sampling basis.

Three types of sensory tests are shown. Nearly everything can be so classified although, in practice, mixed types are sometimes encountered. The first type, preference, is best exemplified in the paired preference and hedonic scale tests so commonly used in research and development work. People are given samples of food to eat, either in pairs, where they are asked to say which sample they prefer, or singly, where they are asked to rate each sample on a scale showing successive degrees of dislike and liking. These laboratory tests and the surveys are alike in that both are concerned with the "affective" response, i.e., how people feel about foods; but, in theory as well as in practice, the objective is different.

Laboratory preference measurements are certainly affected by the same attitudes that determine the questionnaire responses, but we know that preference for foods can vary widely depending on their specific flavor characteristics. The surveys cannot measure such variation; this requires that the foods themselves be present. Ordinarily laboratory tests are not used to establish population levels of preference. They have some value for this purpose, but the results must be interpreted with caution because of the restricted samples of people customarily used and because of the dependence of the results upon the specific conditions of the test. Their greatest value lies in establishing relative preference among comparable treatments. It is assumed that particular test conditions and general attitudes, which seem to contribute most of the variation within the population, will affect all treatments equally. Experience has shown that this is a fairly safe assumption--that relative preference tends to be stable.

Included in the preference type are "quality evaluation" tests. One example would be the typical operation using a small panel where samples are rated on a scale of "poor" to "excellent"; another example might be the work of professional graders of butter, tea, coffee, etc.. Defining such tests as preference will not meet with unanimous agreement--many people would contend that these are primarily difference tests; however, the best evidence shows that, basically, they are evaluating on the preference continuum. They differ from other preference tests only in the scale used and the makeup of the panel.

The objective of difference tests, the next type, is to apprehend flavor differences in order to relate them to processing, formulation, raw materials, or some other factor. The most familiar example is the triangle test, along with its many variants, such as the duo-trio, where we attempt to determine simply whether there is a detectable difference between two test materials. Another example would be the comparison of particular flavor characteristics by means of a rating scale of intensity. These are research and control tools which help us deal with materials and processes in terms of their probable flavor resultants. Basically, we are measuring discrimination, but in a context where conditions are made optimal. Such tests are usually run with special panels--no one worries about the sampling of subjects. This is justified in two ways. First, capacity for discrimination varies a great deal less among people than does preference. Further, difference test results usually would have little meaning as applied directly to the consumer population. They are seldom used to predict acceptance behavior.

Flavor description has been set up as a separate type because, even though it belongs with the difference tests in the sense that it excludes preference and measures only discrimination, it is more complex than the difference tests. The objective is more ambitious--not just to detect flavor changes but to provide integrated qualitative and quantitative information. The most familiar example is called the Flavor Profile, where a special panel works extensively on a given product. This method first determines the flavor properties which are important and develops a naming system. Then the verbal system is used by the panel to analyze the flavor of samples of the same type, by indicating the intensity of each component. Other approaches differ in details, but they are basically the same as the Flavor Profile. Flavor analysis is another research tool whose value lies in the support it can give to the technologist who has to know the relationship between flavor and materials and processes. Again, results require interpretation; no one expects to apply them directly to preference problems.

We have already suggested some reasons why these types of tests may be affected differently, as far as precision and reliability are concerned, by differences among people, environmental conditions, or test situations. Table 2 will help us explore this topic further. It attempts to show the relative importance of various factors which influence the results of sensory tests. Again the surveys are included for comparison. This scheme is proposed only as a means of encouraging controlled thinking on the subject. Certain of its major features have merit, but there are other points that might be hard to defend. For example, the list of factors is far from complete, factors are not mutually exclusive, important interactions have been disregarded, and

finally, the levels of importance vary so much that the variation cannot be adequately represented on the six-point scale. The scale is such a rough one that only the end-points have been tied down, as "0 = negligible" and "5 = very important". Comparisons among factors within tests are probably more valid, but comparisons between tests are supposed to be meaningful, too.

The first three factors listed represent familiar ground. Sensory properties of foods and attitudes toward foods are the things we intend to measure, and it is indeed encouraging to be able to assign high values to them. Sensory properties are high on all of the sensory tests and "0" on the surveys. Long range attitudes are "5" for surveys, "4" for preference tests, and "0" for the others. But now we must recognize two different sets of attitudes--not only long-range response tendencies derived from past experiences which have been well integrated into the individual personality, but also short-range influences derived from one's present relationship to his environment and that in the recent past--such things, for example, as the bad hamburger you ate last week, the raise you didn't get, your knowledge that the boss is very anxious about the newly developed product, or your awareness of the pretty new secretary in the next panel booth. Sometimes one is interested in such short-range effects, but usually they must be considered as only interference because they shift around so much that they have little predictive value. They have most effect on the preference tests, both sensory and survey, but cannot be completely disregarded in the other types because they can also affect discrimination.

The rest of the factors listed, for all practical purposes, may be considered solely as sources of error. To increase accuracy and reliability we have to control them or learn to understand, and account for, the variability they contribute to our test measures.

The next item, physiological factors, covers a broad range--really too broad to permit proper assessment of its relative influence. Here we would include such long-range influences as the need for calories, a person's state of health, adjustments to temperature conditions, age, and many others -- also short-range effects such as the momentary state of hunger, temporary illnesses, or environmental conditions that interfere with the flavor senses. Most of these things can affect sensory preference tests, and perhaps this factor should be accorded a value higher than "2". However, in practice the combined influence is probably not as great as the number of possible causes would seem to suggest. They have been rated a "3" on the other sensory tests because of their potential effect on discrimination, and because discrimination is the sine qua non of these tests.

Basic sensory capacity could properly have been included above, on the grounds that it is physiologically determined, but has been set apart in order to make a special point. That people differ widely in basic capacity to detect tastes, odors, and other stimuli can be easily demonstrated. These differences are due partly to differences in genetic endowment, partly to things that have happened later in life. Concern has often been expressed about the implications of such variations for sensory testing. However, for certain practical reasons, this factor has been relegated to a minor role. The reasons will be discussed later.

The next factor group, motivation and attention, could have been a sub-class under attitudes, but deserves special mention because of its pervasive importance. The quality of any human performance will depend upon the reasons, and the strength of the reasons, which a person has for participating, and the related factor of how closely he applies himself to the task. Therefore, there must be a sufficient degree of motivation for any kind of test, although it must be considered more important where the task is more difficult, i.e., in difference and flavor description. The only reason this factor is not rated "5" across the board is that, in practice, the potential variation from this source can be effectively reduced by use of the right approaches.

Specific test skills are not important for preference testing. They can be designed simply so that almost anyone can understand and respond meaningfully, although it must be admitted that they are not always so constructed. Difference and flavor description tests, in contrast, require people to deal with flavors in ways that lie outside their day-to-day experience, and training is necessary to acquire the new skills.

The factor of language has been included primarily because of its great importance in flavor description. Language is eliminated as a variable in difference tests, and plays only a minor role in preference tests, although attention to the construction of scales and the wording of questionnaires is necessary. But when we begin to deal with flavors specifically and analytically, as we must do in order to get useful results from flavor description, we run into trouble. Communication is in a primitive state. We have lots of words with broad, tenuous meanings. We can say that the test sample smells or tastes "something like" another substance, but we have no ready-made, reliable system for communicating about the flavors as such. One panel member may say "rancid", another "oxidized", and a third "cardboardy", and we will never know whether this represents real differences in perception or just different verbal habits. In consequence, one of the most important aspects of such testing must be the development of a language system and the training of subjects so they will use it reliably.

Does this review point to any conclusions? In surveys, clearly the emphasis should be on getting a broad sampling of the population in order to establish relatively permanent group attitudes, avoiding short-range, "interference" attitudes as much as possible. In all of the sensory tests, the emphasis is on control, but for various reasons. Laboratory preference testing requires that the sensory properties of foods be paramount, yet any measurements will be meaningless unless made within the proper context of permanent attitudes. Hence sampling must be as good as possible within the limiting conditions. Beyond that, emphasis should be on control of the other factors so that all treatments will be affected equally. Then relative preference values will remain fairly constant even though the absolute levels might change under different conditions. In both difference testing and flavor description, the sole emphasis should be on control of conditions and operations to provide for maximum discrimination. This includes the development of test skills, particularly language in flavor description.

With this review as background, let's look at sensory testing operations from another angle, considering now the various things that

may be done to improve their accuracy and reliability rather than the basic factors involved.

1. Physical control of the test environment.

The attempt is usually made to create an optimal environment for sensory testing by eliminating direct sensory interference and reducing sources of distraction to a low level. Control of direct interference with the other flavor senses can be handled in other ways, but elimination of odors is usually accomplished by special laboratory construction and air-conditioning. Laboratory construction helps in the control of other aspects of the environment such as noises, lighting, and temperature which may serve as sensory distractors, thereby reducing attention and lowering performance. Finally, let's mention the importance of laboratory location and design as a help in achieving psychological control. Referring to Table 2, it may be seen that these controls are related mainly to two of the factor groups--physiological, where we attempt to get maximum sensitivity, and the area of motivation and attention, where we want to avoid any interference with discrimination or the more complex aspects of performance.

How important are these physical controls for assuring good results? This will vary to some extent with the type of test, being less important for preference. However, it's the sort of thing you don't question anyway. Control will do no harm even if it is not critical in all cases. Obviously, it would be impossible to sample all of the many conditions under which foods are actually consumed; therefore we use optimal and constant conditions.

2. Psychological control

Psychological control, in the broader sense, is the most critical area. It is here that unpredictable variation can be most troublesome. The human organism is highly responsive and a person's behavior in a given test is a function not only of the stimuli presented and the instructions given, but also is in part controlled by what happened an hour, day, or year ago, by his expectations and desires, and by other unanticipated factors. Discrimination and judgment are not free of such influences, but preference is most affected. Most of the variation is attributable to personality, as long-range attitudes, and it would profit us little to attempt control, because probably we would leave the situation worse than it was before. We assume that their effects will average out among people and be constant from one sample of the population to another.

But certain positive actions can be taken toward control of short-range attitudes which, at the same time, will tend to promote better attention. One step in this direction is the well-known panel booths where each person is shielded from the distracting and "contaminating" effects of observing others. To make this really effective there must also be procedural controls to avoid intentional interference between subjects and by outsiders. A very pernicious kind of short-range attitudes are the biases that can quite easily arise through the subjects' having independent knowledge of the products being tested and, particularly, from having expectations about them, or anxiety about the test

results. The attitudes of management, whether intentionally published or not, can be very potent. Continual safeguards must be used to shield test subjects from such knowledge, or eliminate those panel members who may be affected.

3. Panel selection

Since so much of the unreliability is due to variability in our measuring instruments, i.e., people, and since some people are certainly less variable and more discriminating than others, it might be reasonable to try to improve the situation by selecting the right people. Basically, this is true and the possibility of thus improving accuracy has been demonstrated; but there are many practical limitations. Of course, selection is to be avoided in preference testing, where we want to sample over attitudes, personality, and physiological factors as broadly as possible. The bad guesses that can be made by relying on results from small panels demonstrate this. Selection of subjects does have merit for difference testing and flavor description; however, here we meet the practical difficulties. Selection methods that have been tried include threshold tests for basic sensitivity, ability to identify flavors, ability to repeat quality judgments, and ability to detect flavor differences. But "quickie" tests are of no value. Much work is required to obtain enough data to predict a person's future performance. A selection program usually produces a few good panel members, a few who are definitely inadequate, and a large number in between. Thus, one has gained little. Generally, the best plan is to reject only on the basis of easily detected deficiencies in sensitivity or skill, and to use apparent motivation and interest as the main positive selection factors, since they are rather good predictors of success. One can easily compensate for lack of precision in the individual measurement by increasing the size of the panel, and it is usually more efficient.

4. Physiological control of test subjects.

There is little that can, or, perhaps, need be done about most physiological variables. Sometimes testing is limited to the same period each day to keep the effect of hunger as constant as possible. People tend to respond more favorably on preference tests when they are hungry; also it has been claimed that they are more sensitive to flavors, but the latter is not well established. However, variations in hunger appear to contribute in only a minor way to over-all variability. The major concern here is with maintaining normal sensitivity. People who have colds or are otherwise indisposed are usually eliminated from test panels.

We have already mentioned removal of laboratory odors by air-conditioning. Incidental taste stimuli are not such a problem. Panel members are frequently asked to avoid eating, gum chewing, and smoking for a period prior to testing, although, contrary to what one might expect, there is no good evidence that non-smokers are more discriminating than smokers. People frequently forget instructions; however, this is not critical because most such flavor effects can be eliminated by mouth rinsing and a short rest before testing. Most of the difficulty arises from adaptation and inter-effects among the test samples themselves, which must be controlled through test procedures.

To be realistic and thorough, one more topic must be mentioned. So far we have avoided discussing two things which in practice, probably have most to do with the effectiveness of sensory tests. These are basic test methods and test design. Having a modern, air-conditioned laboratory with panel booths and special lights where well-motivated subjects test under conditions of physiological and psychological control will not assure valid results unless basic methods and procedures are appropriate. They are not only more important than the factors previously discussed but also have a different relationship to the test results. For example, you cannot say that one type of test may be more affected than another. Selection of a method which is inappropriate to the problem at hand may entirely destroy the effectiveness of any test. It is such mistakes, as often as anything else, which tend to create the dissatisfaction with sensory tests.

We do not have time to explore this topic adequately--no more than to give examples of the kind of problems which are involved.

A number of design problems arise because samples are usually tested in groups rather than singly. Inevitably each sample becomes an important part of the frame of reference for the others. This is good because it produces more stable relative measures but it also has undesirable effects. For example, in a multiple sample test, progressively greater adaptation occurs, which can be retarded by mouth rinses and rest pauses but cannot be eliminated. In a preference test on even a short series of samples, there are progressive shifts in attitude which result in lower preference values for the later samples. When samples are presented simultaneously there is often a tendency to prefer the sample in a certain position. One frequently encounters what are called "contrast effects", e.g., an average sample will score high when presented with a poorer one and low when presented with a better one. There is no assured way of eliminating such effects but they can be adjusted for in test design.

The field of test methodology per se is a broad one. Methods are far from being standardized, nor is there any published doctrine for their use. This places responsibility on the experimenter, who must analyze the situation to select a method which is appropriate both to the problem and the test population. Some common faults are: (a) applying a favorite method indiscriminately to all problems, (b) using a difference when the problem is one of preference, or else combining the two types, (c) attempting to infer differences from flavor description data, (d) attempting to infer relative acceptance from ratings of specific properties, and (e) using rating scales which are inappropriate to the product tested or the test population. Knowledge and the exercise of good judgment on the part of those who plan the tests is the only effective means of control in this area of methods and design.

To conclude this discussion, it would be well to return to a point mentioned earlier. The factors we have discussed are not new discoveries; people have always been trying to adjust to them. What we have attempted here is to demonstrate and clarify relationships, a task which must be done in order to place sensory testing on the level of effectiveness which the food industries need.

Table 1. Types of Sensory Tests

	SENSORY TESTS			NON-SENSORY TESTS
	PREFERENCE	DIFFERENCE	FLAVOR DESCRIPTION	PREFERENCE
EXAMPLES	Laboratory preference Quality evaluation	Triangle, Duo-trio, Rating intensity of specific quality	Flavor Profile Flavor "naming"	Questionnaire surveys
STIMULI	Actual Foods	Actual Foods	Actual Foods	Food Names
WHAT IS MEASURED?	Direct response plus attitude	Discrimination (simple)	Discrimination (complex)	Attitudes and opinions
PURPOSE	Establish relative preference	Determine flavor differences	Provide integrated information about flavor	Establish general attitudes - "popularity"
TEST SUBJECTS	Sampling of limited population	Special panels	Special panels	Broad random sample of population

Table 2. Relative Importance* of Various Groups of Factors which Affect the Results of Sensory Tests

Factor Group	Sensory Tests			Non-Sensory (Surveys)
	Preference	Difference	Flavor Description	Preference
Sensory properties (flavor)	4	5	5	0
Attitudes, personality (long range)	4	0	0	5
Attitudes (short range)	3	1	2	3
Physiological	2	3	3	1
Basic sensory capacity	1	2	2	0
Motivation and attention	2	3	3	2
Specific test skills	0	3	4	0
Language	1	0	4	2

* Relative importance is estimated on a 6-point scale extending from 0 = negligible to 5 = very important

References

1. Cairncross, S. E. and Sjöström, L. B. Flavor profiles - a new approach to flavor problems. Food Technology, 6, 308-311, (1950).

2. Peryam, D. R. and Swartz, V. W. Measurement of sensory differences. Food Technology, 6, 390-395 (1950).

3. Peryam, D. R. and Girardot, N. F. Advanced taste-test method, Food Eng., 194, 58-61 (1952).

4. Pilgrim, F. J. The components of food acceptance, Amer. Journ. of Clinical Nutr., (Spring 1957).

5. Pilgrim, F. J. and Wood, K. R. Comparative sensitivity of rating scale and paired comparison methods for measuring consumer preference, Food Technology, 9, 385-387, (1955).

INTEGRATION OF TECHNIQUES IN PROCESS DEVELOPMENT

G. E. P. Box
Princeton University

INTRODUCTION

Quite different problems confront the experimenter at the various stages of development of a process investigation and quite different techniques are appropriate to solve them. As the last speaker in this discussion on the strategy of experimentation it seems appropriate for me to discuss the roles which various statistical methods can play in helping process development along and to try to see how these various techniques fit together.

It is perhaps unnecessary to say that I do not believe in the present existence or possible future existence of any particular statistical procedure which would be best in all circumstances. The current situation in statistics is not that there are a variety of techniques in competition to fill some narrow area of endeavour. Rather the field of interest might be likened to a large unfinished painting. Some outlines have been sketched and a few areas have been painted in but major portions of the canvas are quite blank. We have to recognise where these blank spaces are and to think how they should be filled. Certainly it will do no good to pretend that the painting is really complete, or that none of the parts which have been painted in should ever be retouched, or that all the blank spaces can be adequately filled merely by repeating a pattern which has been used successfully somewhere else.

1.1 Indeterminancy of experimentation

Someone once said that the only time that an experiment can be properly planned is after it has been completed. The more one contemplates this paradox the more one is convinced of its essential truth. Consider for example an investigation involving continuous variables like time, temperature, pressure etc., when the true underlying relationships (represented by the response surface) between the response and the variables is not known. This experimental situation, like most others, is beset by a bewildering number of indeterminancies which exist whether any statistical methods are used or not.

Which variables should be studied? In a given situation one experimenter might regard one set of variables as important and another might include different variables.

By how much should the variables be changed? One experimenter might for instance regard a ten degree change in temperature in a particular context as a large change, while another might think that it was quite small.

In what form should the variables be considered? Should we think in terms of time and temperature or of log time and reciprocal absolute temperature? Or maybe we should not think in terms of these two variables at all but rather in terms of some single function of them.

How complex a model (and hence how elaborate an arrangement of experiments) is necessary in a particular situation? For instance in a given context the approximation that locally the response surface was

planar might be sufficiently close to allow progress to be made using a simple arrangement of experiments based on this approximation. In other contexts the approximation and the corresponding experimental arrangement would be quite inadequate.

We see that in an investigation of this kind we are expected to explore the behaviour of a function of unknown complexity within a space which is not even defined. Seemingly a difficult assignment!

1.2 "You choose the strategy, I will choose the surface".

If more is needed to see how hopeless our task appears, let us play a game in which you decide the strategy for exploring an unknown response surface and I, acting in the capacity of "devil's advocate", decide the surface.

If you choose:	I will choose:
one factor at a time	a surface with large interactions
steepest ascent	a surface with many bumps
*fractional factorial designs	a surface with large 3-factor interactions
second order fit	an exponential type response
grid mapping	a flat plane with a single point sticking out from the surface.

Although for illustration we have confined attention to one particular kind of investigation, indeterminancies which are just as puzzling seem to occur in other situations. One might feel justified in concluding that scientific investigation must be an unrewarding pastime and success a matter of purest luck. We are faced, however, with the embarrassing fact of the phenomenal success of the experimental method over the last three hundred years or so. Perhaps we are not thinking about things the right way. Let us consider the experimental method itself for a moment.

2. THE ITERATIVE NATURE OF EXPERIMENTATION

It is well known that a recurrent theme is found in almost all scientific investigation. This is the successive interplay of the two complementary processes which we shall call processes d and a. The first process, d, is concerned with the movement from conjecture to experiment and the second process, a, with the movement from experiment to conjecture. These complementary processes are used in alternation many times during an investigation. By their repeated employment, the experimenter's knowledge of the system studied becomes steadily greater in the manner illustrated in Fig.1.

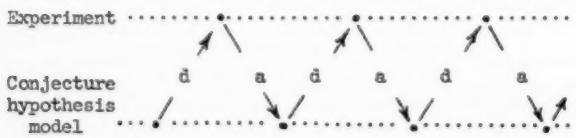


Fig. 1. Iterative experimentation

More specifically:-

process d is concerned with the devising of experiments suggested by the investigator's appreciation of the situation to date and calculated to elucidate it further;

process a is concerned with examination of the results of experiments in the light of all the background knowledge available.

To put it in still another way:

process d is concerned with analysis of hypothesis and synthesis of a suitable experiment

process a is concerned with analysis of the experiment and synthesis of a new hypothesis or hypotheses.

In a statistical context it is customary to relate these processes to the experiment rather than the hypothesis so that process d is associated with what is called the statistical design of the experiment and process a with the statistical analysis. In this paper the words "design" and "analysis" have been avoided until now for, as will be seen from the definitions above and the examples which follow, they are intended to include other things besides the purely statistical concepts usually associated with these labels. With this proviso then the process we are considering is typified by the successive and repeated use of the sequence:

CONJECTURE : DESIGN : EXPERIMENT : ANALYSIS

The iterative situation described above is very different from that implied by certain common usages of statistical methods. For instance "significance testing" as sometimes practiced seems to imply that some irrevocable decision is to be made immediately following the first experiment, no further observations being allowable,and that some dreadful misfortune will befall the experimenter if he errs by drawing a conclusion which is not sufficiently conservative although no penalty will result if he falls into the error of being too conservative. I do not deny that such a situation might occur but I do believe it is unusual.

2.1 An example of classical iterative experimentation

To focus attention on this process of experimental iteration let us consider an example where statistical methods are not involved. We will try to understand the actions and reasoning of some experimenter engaged in speculative research whose object is to prepare a certain chemical C by a route not previously investigated. At the start we suppose he has some, perhaps not very precise, ideas as to the general way that chemical C might be prepared. The iterative sequence would then begin and the thoughts of the experimenter might follow a pattern somewhat like this:

Conjecture: "I believe that in suitable circumstances reactant A would combine with reactant B to form the desired product C.

Design: "From theoretical knowledge, my own experience, and other people's experience of similar reactions cited in the literature, I should think that certain conditions X (defining proportions of A and B, temperature, time etc.)

might be worth trying"

Experiment: The appropriate experiment would then be performed, and this would include a chemical analysis of the product.

Analysis: "After studying the appearance and physical nature of the product and the numerical results, it appears that a little of the desired product C was produced but that a high proportion of unwanted product D was also present."

A single cycle of the sequence has now been completed but the last phase of analysis immediately leads to a new conjecture and the beginning of a new cycle:

Conjecture: "The presence of such a high proportion of unwanted product D could be due to the large amount of water used to dissolve the reactants, which might favour the formation of D".

Design: "I will perform the experiment in the same way as before, but this time I will use a non-aqueous solvent Z instead of water."

He is thus led to perform a further experiment. In this general way his investigation continues.

In some cases this process would be brought to an early stop. Some overriding difficulty might be encountered such as the presence of an embarrassing impurity which would lead the experimenter to abandon this particular route of manufacture. Eventually, however, he would often be led by this method to a mode or modes of chemical preparation which in the laboratory at least were satisfactory.

We see that the scientific method is not, and can never be made, an exact procedure. In particular it is not like an exact method for the calculation of the root of some mathematical equation in which an answer is reached by a unique sequence of steps in calculation. Rather it is an iterative process analogous to the iterative "relaxation" methods of numerical analysis in which faltering (and in the light of subsequent knowledge perhaps even ill advised) steps are made at the beginning of the process. Because the situation is set up so that the investigator has an opportunity to learn as he goes along, however, he can hope to be led along <u>one</u> path to the solution. The path actually followed will differ from one experimenter to another and its length will depend very much on the ability of the person concerned. The objective of the statistician in trying to help the experimenter should be not to try to make the path he follows unique but to supply him with techniques which will help to make the process he follows <u>converge</u> to the correct solution as rapidly as possible.

3. THREE IMPORTANT PHASES OF EXPERIMENTATION

We shall discuss in a moment how the iterative experimental process may proceed in investigations in which statistical techniques are employed; before we do this we consider what sorts of information we might be seeking in such investigations.

To avoid confusing the essential issues let us begin once more by considering a problem where statistical considerations are not essentially involved.

Suppose we were for the first time studying the distance s fallen

by a solid body in a vacuum. Initially one might postulate that s could depend on the levels of a number of variables t, u, v etc. The variable t might represent the time of falling, the variable u the weight of the body, the variable v the volume of the body and so on.

Thus we conceive a possible functional relationship between the response s and the "variables" t, u, v etc.

$$s = f(t, u, v \ldots).$$

Three distinct stages can be considered in the investigation of such a relationship:

i) Study aimed at deciding WHICH variables effect the response. In this example it would be found that (to the degree of exactness here considered) s depended only on t.

ii) Empirical study of HOW the variables affected the response. In this example experiments would enable us to plot a graph showing the relationship between s and t over a certain range of t. We might present this information either in the form of the graph itself or in the handier form obtained by describing the graph by the equation of some simple curve (e.g. an interpolation polynomial) which graduated it in the region concerned. Such an equation would of course make no claim to have any basic meaning and would usually be inapplicable outside the region actually investigated.

iii) Theoretical study which resulted in an equation derived from a postulated mechanism for the phenomenon studied. This might explain more or less adequately 'WHY' the response was affected by the variables in the manner observed. In the present example Newton's gravitational theory leading to the formula $s = \frac{1}{2}gt^2$ would provide such a theoretical equation.

We are thus able to distinguish three component phases of investigation, each of which can involve the use of appropriate statistical methods. These are:

i) Screening studies aimed at delineating the important variables,

ii) Empirical studies aimed at describing certain important features of the relationship between the response and the variables (e.g. the "effects" of the variables in some particular region, or more ambitiously the principal features of the multi-dimensional graph or "response surface" relating the response and the variables in some important region of the variables),

iii) Theoretical studies aimed at discovering plausible mechanisms for the phenomena studied and estimating to satisfactory accuracy the relevant parameters.

All these phases usually involve the same iterative sequence - CONJECTURE, DESIGN, EXPERIMENT, ANALYSIS, so that in terms of complexity although not necessarily of order in time, the route followed resembles Fig.2.

This route should be likened to an unlimited access highway. Few individual investigators do or should proceed from one end of it to the other in any particular study. In some cases an experimenter is concerned only with delineating important variables; he enters at A and leaves at B. In some cases the important variables may have been derived already (or it may be claimed that they are known) and the process

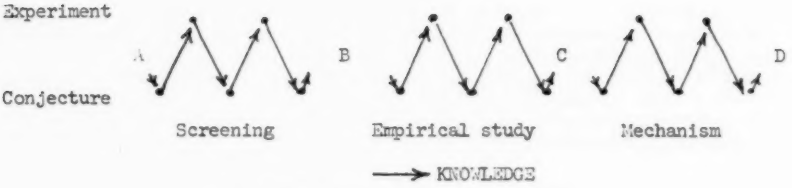


Fig.2 Phases of experimentation

begins at B and perhaps ends at C, or it may go on to D. Or having begun at B it may be necessary to go back to A because it may become clear that the important variables are not known. In differing contexts all points of entry and exit are possible and relevant; "retracing steps" and "doubling back" will frequently occur and the phases we have distinguished may be less distinct. It can be argued that a fourth important phase is the transition from (ii) to (iii). As has been illustrated, for example (1), by careful consideration of the results of an empirical study ideas tending to the development of a theoretical model can develop. We here regard this important process as part of phase (iii).

We shall now briefly survey these various phases of study in the light of the iterative situation we have mentioned and consider what is known of statistical procedures which are appropriate to assist their progress.

4. SCREENING EXPERIMENTS

In the study of the distance s fallen by a solid body in vacuo in time t, we needed to discover that the variables u, v etc. were not affecting s and need not be considered. The situation is illustrated geometrically in Fig.3.

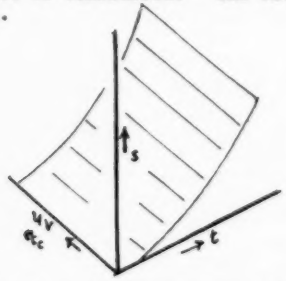


Fig. 3 Redundancy of variables

All the change in s is occurring in the direction of the coordinate axis t. No change is occurring in the direction which represents u, v etc. We can fix the value of u and v (corresponding to taking a section through the model at some point parallel to the axis of t for any values of u and v) and the graph of s against t is unchanged.

Clearly the problem at the screening stage of an investigation is that of probing the space to find those directions in which redundancy occurs. In this example where experimental error is of no importance it would be a simple matter to devise experiments which would do this

at least superficially. Also, for systems as clear cut as this various devices such as dimensional analysis might be employed to help the process along.

Suppose, however, we are confronted with a situation where:

i) very little theoretical knowledge of the system was available;
ii) experimental error was large;
iii) although there was a large number of variables which it was felt might effect the response (perhaps 10 or even more) it was felt unlikely that more than a few of these (perhaps three or less) would be of any real importance;
iv) only a limited experimental effort could be justified;
v) there was no overriding necessity to keep the experimental program simple, and consequently large numbers of variables could be altered simultaneously in the same group of experiments.

Two procedures are known to me, which in some cases at least would seem to be aimed at the solution of this problem. The first usually utilizes fractional factorial designs and the second the random balance designs proposed by Satterthwaite. The details of the latter procedure have yet to be published so I will concentrate on the first.

Bearing in mind characteristics (ii) and (iv) and (v) of the screening situation, it seems we must use a design which, for a given total number of observations, will give maximum possible accuracy for the estimation of the effects of the variables. It is known that this property is enjoyed by designs for which all the variables are included simultaneously and for which there is a high degree of symmetry in the space of the variables. It is enjoyed in particular by factorial designs and fractional factorials where the high precision attained may be thought of as arising from so called "hidden replication". For large numbers of variables it will in general be true that complex relationships which involve all the variables cannot be elucidated without a large number of experiments. For example, where the function can be represented by a polynomial in which terms of order higher than d are omitted, the series will contain $(k + d)!/(k)!(d)!$ constants and we should require at least this number of observations to estimate all of them.

We might proceed, however, by using designs which, although insufficient to estimate separately all the constants if the whole group of ten variables were important, would nevertheless be adequate to estimate all the constants of say first and second order* associated with <u>any particular small group</u> of variables if the remaining variables were unimportant.

Now although the two-level fractional factorials do not completely satisfy this criterion, since estimates of "quadratic effects" are always confounded with each other and with estimates of the mean so that we may wrongly tag as redundant a variable which happens to be

---

FOOTNOTE: * We can of course easily imagine functions which are far too complex to be represented by anything so simple as an equation of second degree even locally but provision (iv) above forecs us to take risks.

passing through a maximum and which does not interact with any other variable, yet this contingency is remote and the designs are in fact of considerable value for the purpose in hand. Their value is still further increased by the addition of centre points. By this means the designs are made to supply at least an "overall" measure of curvature and some rough indication of experimental error.

With fractional factorials, or indeed with any design that can be devised, ambiguities of interpretation can arise. However, usually we will have fairly wide choice of the particular fractional design to be employed and it will be possible so to arrange matters that estimates of main effects and interactions are not associated with other estimates thought to be of importance. The designs also have the great advantage emphasized by Davies and Hay (2) and Daniel (3) that a first fraction can be a "block" in any one of a set of larger fractions, the choice of the second fraction being decided from the results of the first. For a screening investigation employing fractional factorials therefore the iterative process proceeds somewhat along the following lines:

Conjecture: Only three out of ten variables are likely to have appreciable effects.

Design: Use fractional factorial avoiding aliasing main effects with two factor interactions which seem likely to be important and with a view to the possibility of making this design part of a more extensive fractional factorial. Include centre points.

Experiment: Perform appropriate design in random order.

Analysis: Some method of indicating plausibility of various possible explanations of data (such as Cuthbert Daniel's half-normal plot (3). Explanations based on (i) simple redundancy (ii) redundancy after transformation (iii) redundancy after transformation and/or omission of wild observations, particularly considered. General inspection which might suggest ideas not originally in mind of great importance.

Conjecture: Based on those explanations of the data and on those new ideas which were judged "important".

Design: Arranged so that rival plausible conjectures could be separated and "important" explanations further confirmed.

Some remarks on this process may be made. The function of analysis in the scientific process is to facilitate consideration of the data in the light of all possible explanations of it. We shall in this particular stage of experimentation have special interest in detecting which of the variables are redundant. We should not allow this primary condition to impede proper inspection of the data from any other viewpoints that might suggest themselves however. The question of redundancy itself must be broadly viewed. As Daniel has emphasized a fuzzy picture can often be brought into sharper focus by a suitable transformation of the data or by omitting certain of the observations which may have been in error. His "half-normal plot" analysis provides a valuable way of surveying the possibilities. Other methods might be employed. For example an electronic computer might be programmed to calculate values of likelihood or "significance" associated with a very wide range of possibilities (for the problem discussed it might calculate some measure of plausibility for the ten 9-variable redundancies, the forty-five 8-variable redundancies, the 120 7-variable redundancies etc. and the same for various transformations and for different possible

omissions of observations). It might then retain for report all of those having more than a given degree of plausibility.

The question of whether the hypotheses considered in the analysis were in mind when the experiment was planned is of course quite irrelevant in the present context. It is never true that we are barred from deducing new hypotheses from the data for test in later experimentation. If that were so no progress would be possible.

In judging the importance of possible explanations of the data the measure of plausibility is one of the things which must be considered, however at least two other elements must also be taken into account in the planning of further action. First the experimenter must consider whether the effects observed do or do not agree with what he feels to be reasonable. If a hypothesis which seemed contrary to all reason were thrown up as likely the experimenter would rightly require further evidence before he believed it. Second the experimenter must consider the observed effect in relation to its possible utility. If the reality of an effect is in no real doubt but its direction is such that it can be of little value, it will receive less attention than a less plausible effect of potentially higher utility.

4.1 Comments on the formulation of the screening problem

Two comments are appropriate concerning the general formulation of the screening problem:

First, it should be noted that the problem, as stated, is concerned with the detection of redundancy of rather a special kind - namely redundancy of the variables when they are considered in the particular manner defined by the experimenter. In practice where there are say 10 variables which may affect the response we may well find that perhaps half or more of them are redundant in this sense. It will probably turn out however that redundancy of different kinds will occur among the remaining variables.

For example in a problem in which the response "yield" depended only on the ratio of two concentrations the particular decision to consider the separate concentrations as variables and not for example to consider "ratio of concentrations" as one variable and "overall concentration" as the other would mean that no redundancy would be found. To quote a second example, it might be that in a cake-mix only one variable, the acidity of the mix, was really affecting a certain response. If this were so, all variables which affected acidity would be found to have effects, although by making suitable transformations of the variables these effects could be reduced to that of a single variable. Geometrically this means that we are searching the space for redundancy only in those directions parallel to rather arbitrary coordinate axes.

A second point which should be noted is that the appropriateness of the solution given is dependent on the applicability of the assumptions (i) that only a limited experimental effort can be justified and (ii) that a rather complex experimental program can be run.

There are certainly very many practical situations where these assumptions apply, but there are other situations where, provided the experimental designs employed are simple and do not involve too many

of the variables simultaneously, a continuous program of small varieties of process conditions may be installed as part of the normal production routine. This method of "Evolutionary Operation" has been described in (4).

5. EMPIRICAL SURFACE STUDY

The problem at this stage of experimentation is frequently to find optimum process conditions and to describe the behaviour of the response function in the optimal region. The latter part of this sentence is emphasized because knowledge of the position of a single optimum point is seldom satisfactory. This is because (i) knowledge of the local behaviour of the function is essential to allow intelligent operation and control of the process, (ii) to allow recalculation of optimum conditions when external features,such as prices of raw materials, change we must know the local behaviour of the response function near the provisional optimum, (iii) usually the problem involves not one single response but several responses (cost, purity, colour of product,physical form of product etc.). To arrive at a satisfactory compromise it is essential to know the local behaviour of all the response functions in the optimal region.

To allow progress to be made in this problem, iterations of different kinds occur simultaneously:

5.1 Iteration in position.

The results from a group of experiments may be employed to tell us where a further group of experiments should be performed, so that the second group is closer to optimum conditions than the first, or so that it straddles the region of interest more satisfactorily. This is essentially a process of iterating in the position of the experiment in the space of the variables. Examples of experimental procedures which do this are the one factor at a time method, the method of steepest ascent, and methods of surface study which employ canonical analysis of fitted second degree equations. The reader will have no difficulty in distinguishing the components, "conjecture", "design", "experiment","analysis" in such procedures. It will be noted that in practice "steepest ascent" and "canonical analysis" are both examples of dynamic processes of analysis leading to formulation of new experiments.

5.2 Iteration of scales, metrics and transformations

Not only must we iterate on the position at which observations are made in the space of the variables but simultaneously on the way in which we choose to consider that space. Thus initially the experimenter might propose a two-level factorial experiment in which temperature was changed by fifteen degrees and time by three hours, thus implying that at this stage of the investigation at least he regarded these changes as in some way comparable (and consequently that if the response surface were visualized it ought at this stage at least to be visualized so that differences of these magnitudes were represented by equal distances in the space). Results of the first group of experiments might however very well suggest either or both of these scales had been poorly chosen and that they should be made wider or narrower in the next group of experiments.

Again, a group of results might suggest that the metric associated with a given variable was capable of improvement. This is to say they might suggest it would be possible to represent changes associated with that variable more simply and more accurately in terms of some transformed quantity such as its logarithm or reciprocal. It is in fact possible to set out designs which facilitate the discovery of the most adequate metric. As a general rule we would wish to avoid becoming involved in an elaborate form of equation where a simple form would be adequate with modified metrics. If possible, for example, we would try to use a first degree equation with modified metrics rather than a second degree equation. Similarly, we would try to avoid the necessity for cubic equations where second degree equations with modified metrics would suffice. The problem of how to design the experiments and to carry out analysis so as to be led to the most useful metrics is one which is under current study.

Finally, there is the choice of transformation of the variables where these transformations involve more than one variable. It is frequently the case that the response is not most simply expressed in terms of the natural variables such as time, temperature, concentration but in some compound function of them. In the neighbourhood of a maximum the redundancy among the variables studied is usually evidenced by the existence of ridges in the response surface. Study of the response function (for example by canonical analysis) often leads to re-metricization of the problem in terms of a few new variables which are compound functions of the more numerous old variables.

Thus initially we might be working in terms of variables x_1 and x_2 in terms for which the response surface might be a ridge maximum like that shown in Fig. 4a. A 2^2 factorial in this space would be represented by the "square" group of four dots. After a number of iterations

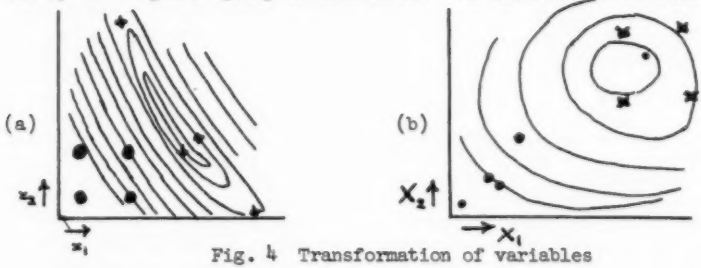


Fig. 4 Transformation of variables

in which the metrics were changed and transformations introduced, we might finally end up with a space defined in terms of new variables $X_1 = f_1(x_1.x_2)$, $X_2 = f_2(x_1.x_2)$, for which the response surface will take on the more symmetrical shape shown in Fig. 4b. A 2^2 factorial in this space is indicated by the crosses. We notice that the "well designed experiment" in the first space becomes a "very poorly designed experiment" in the second space and we see that in fact the concept of "good design" is entirely relative. In any absolute sense an arrangement of experimental points cannot be said to be good except in relation to its power in elucidating the response surface. But since we cannot know what the response surface is like until we have done some experiments we must necessarily follow an iterative procedure in which we try to get a glimpse of the surface and then modify our approach depending

upon what we see. It will not be until towards the end of the iteration that we are in fact getting arrangements of points which are ideally situated for the elucidation of the surface. In the same way, any procedure which we use for analysing the experiments must be based on our best understanding of the situation at the time. Thus, for example, as Dr. Wilson and I (5) emphasized in our original paper, the direction of steepest descent depends on the particular scales, metrics and transformations chosen to describe the space of the variables. At any given stage, the direction of the steepest descent is the best direction of advance in the light of the best information available at the time as to what are the most reasonable scales etc. in which to consider the variables. Like every other process in experimentation it has no absolute validity but is as good as the auxiliary information allows it to be.

5.3. Iteration on the model and the design

Since we usually wish to conserve experiments and to make interpretation as simple as possible, we should begin by making the simplest assumptions we can. At the same time we should choose the experimental arrangement so that a warning will be provided when a more elaborate model (and consequently a more elaborate design) or more replication (hidden or direct) is likely to be needed to make further progress possible. Thus in practice we might begin by performing a simple design such as a two-level factorial or fractional factorial with added centre points, replicating this design, adding other fractions, or augmenting to form higher order designs as appeared necessary.

5.4 Relationship between numerical and exponential iteration

It is of interest to notice how closely the iterative experimental procedure we have discussed matches iterative procedures for the solution of linear and non-linear equations (6). If we have to solve a set of non-linear equations such as

$$f_1(x_1 . x_2) = 0$$
$$f_2(x_1 . x_2) = 0$$

the problem is equivalent to that of finding a stationary value of the function $\Phi(x_1 . x_2)$ where Φ is chosen so that $f_1(x_1 . x_2)$ and $f_2(x_1 . x_2)$ are its partial derivatives. The contours of the function Φ in the space of $x_1 . x_2$ may then very well look like those shown in Fig 4a and iterative methods for finding the solution of the equations are analogous to the methods of experimentation which we have discussed. In particular the one factor at a time method of experimentation is similar to the Gauss process for the solution of the equations, the experimental method of steepest ascent has its corresponding counterpart in the steepest ascent numerical method of Booth (6) and the experimental method of using a locally fitted second degree equation is analogous to Koshal's method for fitting curves by maximum likelihood (9).

6. THEORETICAL MECHANISMS

Either after a number of clues have been provided by results of empirical experimentation such as that described in the last section, or sometimes without such a preliminary investigation, it may be possible for the experimenter to conceive a theoretical mechanism which

might describe the phenomena which are being observed. Thus, in chemistry, kinetic theory may provide such a mechanism. From a mathematical analysis of this hypothetical mechanism it will then usually be possible to produce some theoretical functional form which should relate the observations to the variables. Thus consider a simple chemical example in which a reactant A was decomposed to form B which was then subsequently decomposed to form C. With y_1, y_2 and y_3 representing the yields of A, B and C at time t under certain well-defined assumptions the following set of differential equations represent the system:

$$(1) \qquad \frac{dy_1}{dt} = -k_1 y_1$$

$$\frac{dy_2}{dt} = k_1 y_1 - k_2 y_2$$

$$\frac{dy_3}{dt} = k_2 y_2$$

In this particular example(although not usually)an explicit solution of the differential equations is available, for example, the yield y_2 of the intermediate product B at any given time t is given by

$$(2) \qquad y_2 = \frac{k_1}{k_1 - k_2}(e^{-k_2 t} - e^{-k_1 t}).$$

Suppose that the results of a number of experiments are available. In this case for example y_2 would have been observed at various values of t. The problems then are (i) to discover whether the assumed functional form is adequate to describe the data, and if it is not in what way it is not; (ii) if the functional form does fit the data to estimate values of the unknown parameters (i.e. the constants k_1 and k_2 in our example) and to determine their precision as measured by their standard errors and confidence region. Recent work undertaken on this important problem (7,8) has shown that by using the numerical methods it is possible to set up a general program on an electronic computer which will, by an iterative process analogous to the experimental process we have discussed, answer all the questions posed above and furthermore that this can be done whether the functional form is given explicitly as in equation (2) above or implicitly as for example by the differential equations(1).

It is found that such a program can be of great value in the formulation of theoretical mechanisms. The experimenter may start off with a theoretical model which he feels is almost certain to be inadequate but which will at least crudely describe the process involved. Analysis of the "residuals", that is to say the discrepancies between the observed and calculated values when the best possible fit has been obtained, then enables him to see in what way his first attempt at a model needs to be modified. With the model suitably altered a second set of calculations will then be carried out on a computer. If this model is again found inadequate the nature of the inadequacy is studied and this again leads to further modification of the model. In this way finally one or more plausible mechanisms may be found which adequately describe the data already available. It will be realised that such a process need not lead to a correct model but only to a formulation which is descriptive of the data already collected. It is now necessary to set the computer a new task, that of designing experiments

which most severely test the model so far postulated. We thus set up an iteration of successive analysis and design which together can lead to convergence on the correct model.

7. CONCLUSIONS

What morals, if any, can we draw from this discussion? Some of these are as follows:

i) Different situations require different techniques and major attribute of a good statistician should be to recognise the different situations and to have available the different techniques.

ii) Much experimentation is iterative and it is the statistician's function to assist this process. The iterative nature of the process must affect his attitude both to design and analysis, and to such specific things as significance tests.

iii) A strategy of experimentation for continuous variables is only good or bad in relation to the surface which is explored. Thus, as has been so well brought out in the paper by McArthur and Heigl (10), performance on the black box assesses a strategy not in relation to its absolute value but in relation to the surface which is in the box. Because of this the important question we must ask is what sort of surfaces occur in our particular type of work? When we know this we can devise a strategy which will be appropriate. The particular techniques of experimentation which I devised in cooperation with Dr. Wilson and others were themselves very much a product of the iterative process. We tried something and if it seemed to work we kept it and if it did not we rejected it or modified it as seemed appropriate. The emphasis which the second order part of this strategy places on canonical analysis etc. is symptomatic of the fact that the major characteristic of the majority of surfaces which we studied was the existence of ridges or, to express it in a different way, the occurrence of redundancy or partial redundancy not in the variables as studied but in functions of these variables. Such experience as I have had in other areas has led me to believe that this phenomenon may be fairly common. The important thing is clearly to make a study of the types of surface which occur in practice. So far as chemical problems are concerned, this we have tried to do over the past six or seven years in two ways: firstly, by noting the types of surface which have arisen in as far as they can be elucidated by the techniques we have used; and secondly, by studying the characteristics of theoretical surfaces. With the available modern computing equipment with which differential equations can be readily integrated using numerical methods, the elucidation of a wide variety of surface types can be accomplished.

iv) A point which we have not been able to discuss fully in the body of the paper is that in practice few problems of optimisation are concerned with a single response. Usually there are a number of responses such as yield, purity, colour, physical form etc. and, although in theory we could combine all these into a single criterion of value or cost, this is in practice not a useable procedure, for it requires the equating of incommensurable quantities. Questions like: How much does a bad smell cost? At what cost should we assess the blowing up of the plant? How often can we

kill all the fish in the river? etc. are not easy to answer. A practical solution of the problem which we have adopted is that of maximising the principal response such as cost subject to a series of restrictions necessary to keep the other responses at satisfactory levels. A procedure analogous to linear programming is then employed.

v) In the response surface problem, what is usually needed is not just the position of the maximum but also the nature of the function in this neighbourhood. The existence of insensitivity to changes in specific direction (i.e. ridges) often make it possible to find manufacturing conditions which are near optimal for several responses simultaneously.

vi) A point which must not be lost sight of is the critical importance that an experimenter's knowledge of the problem plays in any experimental process. I have always felt that the statistician has at least as much and possibly more to learn from the good experimenter than the experimenter has from the statistician. In the process of investigation we have described, the experimenter's mind must be in the circuit of iteration particularly at the analysis and conjecture stages. This means that the statistician's analysis must be so expressed as to be clearly comprehensible to the experimenter and it must be such as so far as possible to present all the facts in an undistorted way. The statistician must remember that it is quite possible that the experimenter will see things in the data which he cannot see because of his lack of special experience in the field concerned. Therefore, he must not filter pieces of information which he feels are irrelevant but which in fact may be critical. The experimenter must be given opportunity to assess the data in relation to hypotheses other than those which are in the statistician's mind.

vii) The process we have described is rather like the operation of a servomechanism and the rules for good operation are similar to the rules for good operation of such a mechanism. For example, for efficient operation of such mechanism the signal to noise ratio must be reasonably high, otherwise the system will be unstable. Considerations of the plausibility of various hypotheses should be thought of in this regard rather than in terms of formal significance tests.

References

(1) G.E.P. Box and P.V. Youle, "The exploration and Exploitation of Response Surfaces: An Example of the Link between the Fitted Surface and the Basic Mechanism of the System", Biometrics, Vol 11, No. 3 (1955).

(2) O.L. Davies and W.A. Hay, "The Construction and Uses of Fractional Factorial Designs in Industrial Research", Biometrics, 6 (1950).

(3) Cuthbert Daniel, "Fractional Replication in Industrial Research", Proc. of 3rd Berkeley Symposium on Mathematical Statistica and Probability (1955).

(4) G.E.P. Box, "Evolutionary Operation", Applied Statistics (1957).

(5) G.E.P. Box and K.W. Wilson, "On the Experimental Attainment of Optimum Conditions", J. Royal Statistical Soc., Series B (1951)

(6) A.D. Booth, "Numerical Methods", Butterworths Scientific Publications (1955).

(7) G.E.P. Box and G.A. Coutie, "Application of Digital Computers in the Exploration of Functional Relationships", Proceedings of the Institution of Electrical Engineers, 103 Part B, Supplement No.1 (1956).

(8) G.E.P. Box, "Notes on Non-linear Estimation", Statistical Techniques Group Report, Princeton University.

(9) Koshal, R.S., "Application of the Method of Maximum Likelihood to the Improvement of Curves Fitted by the Method of Moments", J. Royal Statistical Soc. 96 (1933).

(10) D.S. McArthur and J.J. Heigl, "Strategy in Research", Esso Research and Engineering Company (1957).

STATISTICS, QUALITY, AND CONTROL; THE ESSENTIAL TRIO FOR MODERN COMPETITIVE PRODUCTION

Russell V. Johnson
The Johnson Gage Company

On an occasion such as this, when we are marking the Twelfth year of the American Society for Quality Control, we might be given good cause to reflect on the purposes and, more than that, the achievements of SQC as applies directly to each and every one of us who believe in and practice this science. There are those among us who can point to successful endeavors in applying an SQC program within our plants, to real results of economic value and profit to our organizations, and to genuine contributions to our constantly advancing industrial way of life. But, undoubtedly, there are also among us those who have been frustrated in attempts to apply statistical means to better production rates, lower costs, and increased quality of product. And, it is probably correct to state that there are a dangerous number in the latter category who, for one reason or other, have been unable to effectively establish an SQC program in their own plants.

Far be it from the writer to imply that there is a simple solution to the many problems involved in getting an SQC program to work. But it appears that in a great many instances we have forgotten, or at least neglected, the intended practical aspects of this modern production tool; that we have become lost in the maze of charts and graphs, thwarted by management's misunderstanding of our intents, and befuddled by the political aspects of "..it was good enough thirty years ago, so why try to make it better today ?"

It is not the intent of this paper to attempt to remedy all of the problems encountered in trying to get a sound, sensible SQC program initiated in any of your plants, but perhaps if we will consider the basic fundamentals we will better arm ourselves toward getting the most out of Statistics, Quality, and Control.

It is the humble opinion of the writer that an improper balance of power exists in most plants where SQC is only in attempted operation. It is his opinion that the unequal or, should we say, the unrealistic attention, either excessive or inadequate, upon any of the three basic principles of SQC is one of the main causes for the unsuccessful application of an SQC program.

The main theme of this paper will deal with the coupling of effective instrumentation with practical statistics as an absolute necessity. Because of the writer's personal background concerned with the field of screw thread measuring devices, much of this paper will concern itself in that field. The various means by which a screw thread may be measured will be discussed purely from an objective standpoint, and it will be pointed out that effective instrumentation is not only an integral but a mandatory factor in any Statistical Quality Control operation. Finally, it is the opinion of the writer, and others in his field, that proper instrumentation has been sorely neglected by those who are directly concerned with the statistical control of product quality to the degree where-by this entire movement in modern production has been seriously endangered.

The writer does not claim to be a statistician, although he does have a fair working knowledge of most of the currently popular mathematical bases used in SQC today. Being a member of the gage-making industry, it is only fair to state that it is none of our business to try to state whether or not your product is acceptable from a quality standard, we are solely charge with designing equipment which will tell you just what your product actually is from a qualitative standpoint. There then remains the factor of control which is, we feel, directly related to the coupling of proper and effective instrumentation with proper statistics.

Let us, for the sake of clarification, review the three basic divisions as contained in the very title of this paper, namely, Statistics, Quality, and Control.

For many years before these words were joined as in the name of a process or an organization, industry was attempting this concept. It was shortly after World War I when a statistical application to inspection methods was first started but this specific title was lost for many years, even though it was practiced. Since approximately 1940, there have been upwards of fifty different statistical approaches and theories expounded. In fact, if one will note the articles published here for this Eleventh National Convention, there are even newer approaches through applied new statistical concepts. Actually, each type of industry, Chemical, Textile, Rubber, Metal-working, Food Packing, has its own unique statistical approach toward the control of product quality. There are charts, diagrams, distribution curves, all too numerous upon which to elaborate, and there are most likely at least a half dozen in each group to meet the needs of every industry today. We do not infer that we have yet come upon the most fool-proof, the most practical, or the simplest statistical devices, for we must strive to better each system, to determine its flaws, and to ascertain if it is too cumbersome to effectively operate in our own situation. We do not make the inferrence that there are too many statisticians in the world today, but it is the writer's personal knowledge that there are plants today where SQC is no more than a group of very well compiled charts and data sheets, where the Quality Control Department is like a department of higher mathematics, where for certain (sometimes personal) reasons the director of the SQC program occupies a chair in the proverbial Ivory Tower. Nobody can criticise him because nobody can understand his mathematical genius. This is a harsh statement to make, nevertheless it is known to be true in certain instances. Summarily, although we realize that there is a definite need for Statistics to simplify procedures and to corroborate more complete data, there cannot be placed too much emphasis on statistics alone. It is not the intent of this paper to place discredit upon those whose profession it is to devise these mathematical concepts which can be practically applied to facilitate production and inspection, in fact, that would be as far from the truth as anything could be, but we do wish to point out that UNLESS we can effectively couple a sound statistical basis with effective instrumentation we will not meet with managements requirements for producing better product at lesser cost. It cannot be done with statistics alone.

The direct result and most intangible aspect in SQC is Quality itself. To borrow a mathematical term perhaps it would be safe to say that quality itself if the biggest variable. There are too many

factors which influence the quality of the final product. First of all, we are dependent upon our machine or manufacturing capabilities because statistics and instrumentation will not make parts alone, the machine must make them. Obviously, one of the first steps to consider is the capability of our manufacturing processes and our inspection equipment. From there, we must ascertain the quality level we must maintain in order to meet or better competition from a quality standpoint. Doubtless, most sales department complaints hinge on inferior quality, or on a delayed shipment caused by rejected material. Outside influences such as customer requirements, and competitive quality and price, cause us to attempt to set a level. These external factors and other influences are not to be discussed in a paper of this sort, but we must realize that once a quality level is set, whether through external or internal means, or both, it then becomes the task of applied statistics and effective instrumentation to garantee that we will meet and hold these desired and demanded levels.

This brings us to the subject of gaging and/or instrumentation as it must be applied in an SQC program if that program is to reap its full benefits.

First of all, we must define the purpose of gaging. We might first define a gage, what it is, and what is its intended purpose. One more popular definition of a gage is that "it is a device for determining accuracy and inter-changeability." However, there is an implication in this definition that a gage is not and should not be considered as having any aspects of control over the part being produced. Obviously, we must not allow ourselves, in this modern day, to interpret or define a gage as being simply a judge of whether or not a part is good or bad, acceptable or rejectable. We do not wish to make unfair or unreal statements or suggestions as to the effectiveness of a final inspection operation either, for it is an integral part of the quality determining factor of items about to be shipped from the manufacturing plant. But now that we have gone through an era where we have come to realize that quality cannot be inspected into a part, it must be built into it, a gage then becomes more than a device with which to determine the acceptability of final product. A gage now becomes "a device to facilitate production and garantee the quality of the items produced." Thus, when used in this latter defined position, a gage is no longer something which merely inspects, it becomes an integral part of the production apparatus. A gage becomes a tool, homogenous with cutters, grinding wheels, etc., and it is only when a gage is regarded as a production-used tool that its fullest measure of effectiveness can be realized. We are not attempting to imply that a gage should be used in the same manner as most tools are used, particularly from the psychological standpoint where tool adjustment normally means variation in product dimension. Rather, a gage must be used for the puspose for which it was designed, often with care and always with respect. A piece must be free from foreign matter to a reasonable degree when it is inspected, and the gage must never be forced to fit the part. The gage becomes the machine operator's second brain, his sixth sense, helping him to judge when his equipment must be re-set or, sometimes, when the operation must be stopped altogether. It is still the unfortunate condition in many plants today where both management and operators regard gages as negative implements, policemen about to make an arrest in a punitive fashion, penalizers, and deterrents to high production. A gage, conversely, is far from that, when it is a correct

gage from the standpoint of design and cost. Correctly used, it is the operator's biggest aid in helping him to keep his machine running at all times so that he can get a maximum of acceptable product output with a minimum of down time and rejects. Unfortunately there are instances where management is guilty in opposing effective gaging practices because they do not realize that gages, correctly designed and used, will increase (not decrease) their output. Then too, management often times has had experience with unreasonably expensive inspection procedures and has returned to antiquated and supposedly less costly inspection devices which are actually the most costly items which could be used. Regarding the cost of inspection, management often forgets that it is not simply the cost of the gage itself but the cost of the inspection operation expressed in terms of (a) time for the inspection operation, (b) wear factor on both operator and gage alike, and then (c) the cost of the gage itself. Obviously, it is the sum total of these three factors which must be considered in effectively selecting inspection equipment.

With specific regard to the manufacture of threaded product and to the gaging of screw threads, how can effective instrumentation ally itself with statistics to assure the quality level of manufactured product ?

We must bear in mind that the ultimate in a good statistical plan is to inspect a minimum number of pieces and to determine, by the results of the inspection process, the quality level of the items being produced. Obviously, there are natural defects within any sampling plan. True, much has been done to eliminate this chance factor, but on a purely mathematical basis there are still areas in which, through natural mathematical law, complete accuracy cannot be predicted or expected. Taking this fact then, and recognizing that there are limitations inherent in even the best of statistical sampling plans, we must minimize the effect of faulty or incomplete inspection methods, of gage error and of observational error, since these would only tend to multiply the error in the system. It is axiomatic to state that the fewer pieces inspected (and this is the ultimate of an SQC program) the more critically they must be inspected. An example of this, although not too practical, might be a run of screws from a particular machine to which we wished to plot an $\overline{X}$ and R Chart. Let us further assume that we were going to divide our chart into increments of .0001" graduations. If we were to measure the pitch diameter of these screws on a dial indicator type comparator, would it be correct to use a dial which was graduated in .0005" increments ? Obviously not, in fact, when we couple this effect we would be very lucky to wind up with any kind of authoritative chart; it would probably resemble a proposed trans-continental twenty lane highway laid out with a very dull pencil.

The above cited example is in no way intended to give an over-simplified answer to the problem of correcting all approaches to the measuring of a screw thread. Yet, we feel that there are many instances, not only in the field of screw thread measurement but in other fields as well, where the basic fundamentals of instrumentation are so thoroughly incompatible with control chart requirements as to render the actual inspection operation completely ineffective. Dealing specifically with threaded product, we will discuss the various means by which a screw thread might be inspected and point out the merits and limitations in the various basic types of gages used today. We felt that we

could impartially discuss these various types of gages and systems for over the past fifty years practically every major development in the field of screw thread checking has come from our company.

Generally regarding externally threaded product, the Go thread ring gage is usually given as the criteria for assemble-ability checking of a male threaded part. This gage is actually a functional, or cumulative check which theoretically assures that the combination of errors in the thread, the lead, pitch diameter, angle, out of round, taper, etc., is such that the part will assemble with its mating female thread which would be effectively no smaller than the Go thread ring gage. Purely from the standpoint of assemble-ability, the correctly made and unworn solid Go thread ring gage serves this purpose fully. We are in complete accord with the gaging philosophy that all other devices for measuring screw threads cumulatively must eventually square and agree with the solid Go thread ring gage. Yet, there are certain factors or defects in manufacturing processes which can never be detected through the use of a Go thread ring gage. What about a condition of out of round, of off lead, of back (or front) taper ? Obviously, a fixed gage, such as a Go thread ring gage, will not reveal these defects. We must bear in mind then that the Go ring gage can do nothing more than to tell us the maximum effective size of the part, its assemble-ability, and should be regarded solely as such. The prime limitation of the Go thread ring gage, aside from the fact that it is merely an assemble-ability check, is its high cost of use from the standpoint of time and wear. It is extremely un-economical and impractical for use in a high production area.

A second type of Go thread gage is the Roll Snap Gage, pioneered and first manufactured by our company in the early 1920's. This gage consists of a pair of annular ribbed rolls mounted on eccentric pins and is an extremely economical gage to use due to its long life, fast action, and liberal adjustment. However, from a purely technical aspect this gage is quite inadequate as it gives a diametral check in one plane only for pitch diameter, lead, and angle. It is a linear contact gage only, and is incapable of detecting thread drunkeness, a poor starting thread, and sometimes out of round, as it only engages the part being inspected at two points, 180° apart from each other, and in one plane. Furthermore, unless the float or axial movement of the Go rolls is restricted it would rarely detect off lead or irregular helix. A Roll snap gage should never be used without a simultaneous spot-check with a Go thread ring gage.

Another type of Go gage in use today consists of two opposing flat anvils upon which a straight line thread form has been generated. This type of gage shares the limitation of the Roll Snap Thread Gage in the fact that the gaging elements are located in one plane only. This gage does not wear as well as the Roll Thread Snap Gage, however.

A third type of "fixed" gage is the Ring-Snap Gage which embodies the "snap" action principle of the two above mentioned gages, but which consists of threaded segments which engage the part peripherally, such as a ring gage does. Technically, this gage shares the features of the ring thread gage with the economics of the Roll Snap Gage. The gaging members are mounted on eccentric studs for easy adjustment and the segments out-wear ring gages many times over.

The aforementioned gages constitute the more popular present day means by which external screw threads are measured on an attribute basis for assemble-ability. As previously stated, fixed (Go) type gaging is generally regarded as the final criteria for acceptance or rejection of product at maximum assemble-ability limits. But while we are concerning our remarks to "Go" type gaging let us, for a moment, consider the application of variables type gaging for cumulative (Go) checking of an external thread. It would be well to bear in mind at this time that the purpose of "Go" gaging is to determine assemble-ability. But in this modern day we have come to realize that the mere fact that a part will "assemble" does not necessarily mean that the part is good enough. One example of this would be a condition of out of round or taper where the maximum condition would not exceed the Go gage but the minimum condition or size might well be below the minimum limit. In actual assembly, a part so constituted would assemble, but would it remain assembled when subjected to conditions of stress or vibration ? The entire load would be concentrated on the area of maximum size and the chances of the part remaining assembled are very slim. By using variables type equipment these conditions would be readily detected.

To discuss briefly the various types of indicating (mechanical) Go thread gaging equipment being marketed today, there are basically only two types, namely Roll and Segment Comparators. These categories can be divided into two and three point contact on roll comparators, while the segment type comparator is essentially a two point gage unless the size of the part being checked is equal to the size of the gaging elements.

The roll type comparator, which engages the part on a linear basis in general is not as critical as the peripheral contact of a segment type comparator. A two roll comparator will not always detect irregular helix because it has the same limitation of a roll snap gage, as the part being gaged is gaged only in one plane. Ovality can be detected in a two roll indicating thread comparator provided that the piece is rotated, but this is generally difficult to do since this usually means that the part must be re-inserted in the comparator and a fly-reading taken. A three (or tri) roll comparator, which engages the piece at three places, usually 120° apart, will detect triangulation, but not pure ovality. The three roll comparator is more apt to detect helical irregularity than a two roll comparator, but because of the line-contact features in both comparators there are certain of the helical distortions which could remain un-detected, also a poor starting thread could conceivably be missed using a roll type comparator. However, regardless of these slight technical limitations, roll type comparators enjoy high popularity primarily because of their long life of gaging elements and, in some instances, the effects of dirt and other foreign matter are reduced.

However, we must remember that a Go gage or comparator should duplicate, as closely as possible, the mating part. For this reason, then, a segment type gage or comparator is more critical because of the peripheral or circumferential engagement on the part being inspected.

Up to this point we have concerned ourselves primarily with the various means by which a male screw thread might be checked for assemble-ability. We have also suggested that by using a variables type gage we will not only determine whether or not a part will assemble

with its mating part, but we can further detect certain errors through the use of variables type gaging which cannot be detected through the use of fixed gaging methods. Also, we have suggested that through the use of indicating type gaging a machine operator can determine trends toward maximum or minimum limits, thus assisting him for machine and setup control. But what about "Not-Go" gaging ? Why should we gage for not go, and, in the case of a screw thread, how can we best do this?

First of all, we must recognize that a common screw thread differs from a plain cylindrical part in that one cannot use an effective (or cumulative) check at both the maximum and minimum limits. We have to check the various elements in the thread to assure that a part will not only assemble but will remain assembled. We must also recognize that the tolerances expressed for a screw thread are cumulative, that is, the tolerances for pitch diameter are cumulative and include the variations in lead and angle. Thus, the full pitch diameter tolerance cannot be used unless the lead and angle of the thread are perfect. Therefore, it becomes necessary to check the actual pitch diameter of the product in addition to the assemble-ability or effective size. We must bear in mind that the pitch diameter is actually the amount of metal left in the screw, and an error in lead will give an effectively larger reading on an actually under-sized screw. An error of .001" in lead will give an effective pitch diameter reading which is .00173 (or nearly .002") larger than the screw actually is. Further, one can determine by the process of elimination whether or not the elements of lead and angle are within tolerance by contrasting the effective size reading with the actual pitch diameter of the part. It is possible to check the other main elements of the thread, namely lead and angle, by other means, but for all practical purposes this is usually inconvenient. The simplest element to check is the pitch diameter. Thus, the two reasons for measuring pitch diameter are (a) to determine if the thread is within prescribed limits from a pitch diameter specification, and (b) because the pitch diameter is the simplest element to check.

How should the pitch diameter of a male thread be measured ? Which methods are currently used and which of these means are correct ?

To recall our definition of not-go gaging for threads, we are reminded that a not go gage should check one element of the thread only, namely, the pitch diameter. One of the most common, and certainly the most incorrect, ways of checking pitch diameter at the minimum limit is by the use of a not go ring thread gage. There is absolutely no validity in the use of this gage since it cannot measure the pitch diameter of the threaded part. An off-lead condition in the part would cause the part to bind or be kept out of the not go ring. Therefore, when this condition occurs, one would assume that the pitch diameter were not undersized. Furthermore, a bruised or "fat" starting thread would prohibit the part from entering the not-go ring gage, thereby making it an acceptable part. Also, as previously mentioned, back taper or out of round would keep the part from entering the not go ring. Last, and perhaps most important, the theoretical principle of "gaging by NOT gaging" cannot be justified. Actually, a not go ring gage conceals, rather than detects, error in the thread. It gages the thread by not gaging it. There is no technical validity in this concept of gaging.

In order to correctly measure the pitch diameter of a thread, the effects of lead and angle must be eliminated. The cone and vee roll is

the most suitable and correct means and duplicates the principle of the pitch micrometer. This can be achieved by either a pair of rolls, one cone and one vee, diametrically opposed on either a fixed or indicating gage, or by the combination of three rolls, one vee and two cones, on an indicating type comparator. The former principle is embodied in the not-go portion of a roll snap gage and in the aforementioned ring snap gage. For this reason a roll snap gage is correct at the not-go section, but only when the not-go rolls are a cone and vee. We must bear in mind the fact that we are attempting to determine the amount of metal remaining in the screw at the pitch line, and not the amount of metal removed from the blank at the pitch line. A radius roll, or rolls, or a thread wire actually measures the removed metal (or the void across the thread) while the cone and vee measures the amount remaining at the pitch line. Further examination reveals that thread wires should not be used on product for the reason that slight surface distortions along the flank angle of the thread can cause the wire to seat in a "low" spot and give an erroneous reading when used with a micrometer. Thread wires should only be used on plug gages or taps or threaded product where lead and angle and surface finish are well controlled elements.

Although the previous paragraphs have discussed briefly some of the various means by which an external thread may be checked, it is not the writer's intent to imply that these constitute all known methods. These cited types of gages and comparators are the more commonly used today. It is our basic contention that all mechanical devices used for the measuring of male screw threads, in order to be technically correct, should emminate from the principle of the Go thread ring gage and the Pitch micrometer. The purpose sought in discussing these various means is to point out the technical and economic limitations and advantages in each type of gage so that those who are concerned with implementing their SQC programs, relative to the production or inspection of screw threads, with correct and effective instrumentation will have a better knowledge of which type, or types, of equipment to use.

It cannot be stressed too vehemently that a control chart is absolutely no more valid than the means by which the data for it is compiled, and that hinges directly on the selection and use of effective instrumentation. If a gaging system which conceals error is used, then a very false picture of a high quality level will be the result. This condition cannot be tolerated if one is to advantageously apply SQC. Your customer goes not by the appearance of the control chart, but by the actual acceptability of the purchased items. The inspection department is the customer's representative and management's guardian, and it is only as efficient as the inspection equipment used.

Finally, we have observed that quality can be only the result of correctly applied instrumentation. When we do this on a statistical basis then, and then only, do we have the statistical control of product quality. One can have simple quality control without ever seeing a control chart, but that is not the intent or purpose of SQC. It is much more of a challenge for a statistician to be sympathetic to the practical demands of the manufacturing or inspection departments than it is for a machinist or an inspector to be able to comprehend the various axioms of frequency distribution, probabilities, etc. Many clinics have been run on an in-shop and local level to acquaint operators with the theories of SQC, and with success. Perhaps more should be organized for the one purpose of acquainting SQC people with modern instrumentation means.

Without statistics one merely has a form of control; without effective instrumentation, quality becomes no more than an imaginative commodity on a chart basis.

It may well be that the difference between the success and failure of an SQC program could lie in the fact that too much attention was given to statistics and not enough thought was given to effective and proper instrumentation. Most assuredly, the converse might also be true. We must recognize that the coupling of effective instrumentation and a sound statistical basis are the two basic requirements in a true statistical quality control operation. Gages which conceal rather than reveal defects are of no use, they give false data for charts and poor quality of shipments. When, and only when, we give adequate and realistic attention to both phases in this program, then we have SQC.

A RANDOM SAMPLE OF CONTROL PHILOSOPHY

Karl J. Bauer
Abbott Laboratories

The paper will trace control procedures and philosophy from the pre-Statistical Quality Control era thru the maze of charts and data to the present simplified methods actually in use. Particular emphasis is placed on simplicity, training and a program of Quality Attitude which attain the desired results.

QUALITY CONTROL'S OBLIGATION TO MEASURE TERMINAL QUALITY

Captain Ross F. Mahachek, US Navy
Bureau of Aeronautics, Navy Department

For military aircraft and equipment, it is especially true that attainment of a high measure of quality at the time of acceptance is not an end in itself. It is only a step toward the prime objective of effectiveness of the item in military use. On that basis we could now ask ourselves a pertinent question: Are we concentrating on control of quality based strongly on factory and laboratory experience while achieving only slow development of the feedback of consumer experience?

Such an unbalance could easily occur. Quality measurements up to the factory door are relatively easier, handier, more orderly, more satisfying, and better developed. In contrast, feedback of consumer experience, from events scattered around the military world and scattered along the calendar, is difficult to obtain and is laced with negligible cases and guesswork. It has been slow in development and has been generally frustrating. We talk much more about our factory experience and calculations than about the user's experience. We tend to measure our successes by using factory measurements and tests. These are more firm and immediate than the information on ultimate effectiveness which we receive from the military consumer months later and involving widely varying environments.

We receive only a part of the vital statistics concerning the quality and reliability of our equipment in the consumption phase of its life. Except for the limited areas covered by Field Service Representatives it is only when the faults are fairly severe or directly troublesome that unsatisfactory reports will be volunteered by the users.

Let us consider the numerous lesser cases of faulty manufacturing or inspection. They tend to go undiscovered or to go unreported if they are discovered. Even when feedback reports on them do come in, they are often hard to handle in scattered small quantities. Some are ambiguous in phrasing. A considerable number are obviously the work of persons trained only in other fields. Some are even slanted toward assuming faulty manufacturing or design weakness when faulty handling by the user is actually involved. Only when a fairly full coverage is received will it be practicable to discern a pattern upon which correction of non-critical faults can be justified. The manufacturer thus finds little basis upon which to make improvements or corrections of these unproven lesser faults. The contracting activity likewise does not have enough basis to drive for improvement of an apparently nebulous condition.

Why worry about these lesser faults which have not been forcing us to action? Why go looking for minor troubles when there are major flight safety troubles to work on? The answer is that the total of these numerous lesser faults may well accumulate to an appreciable total equivalent of consumed dollars or lessened effectiveness. This loss might be reduced by investing a much smaller sum in mass methods of obtaining more detailed discovery, reporting and corrective action. Furthermore, the dominant goal of effectiveness at the front line may well be helped through a gradual concurrent rise in basic reliability of the article. The change in cost or in reliability will not be great in the average case, but the number of profitable cases may well be very large. TWO percent gain could be considered small but two percent of $5 billion is

still a $100 million which justifies a considerable effort. A modest percent improvement in economy or reliability from increased cultivation of this concealed area will seem more promising when we consider that our initial lack is only of information and practicable mass communication and not lack of technical ability to take profitable action thereafter.

Now let us briefly examine some of the action already taken or underway. First of all, we are not talking of a new problem. Action has been underway for decades. It was only stunted in its results by the fact that the information tended to remain stubbornly concealed in the grass roots. The cost of releasing it, by methods then known, seemed higher than the potential savings.

The two valuable tools recently applied to these field and factory reports of faults are the electric punch card and statistical methods. A number of airframe and aircraft equipment contractors are among those manufacturers who have been applying the punch card for several years with varying but encouraging degrees of success. As an example, one of the successful users, an aviation manufacturer, recently reported excellent results in his 3rd and 4th years of such application. However, his initial 2 years of application had included a characteristic period of deep discouragement arising from the unforeseen volume of reports which overwhelmed the original set-up and from the usual temporary skepticism of many personnel whose cooperation was needed but could be won only after a demonstration of some successful results.

The Navy Bureau of Aeronautics has, for over a year, been applying electric punch cards to a new simplified "Failure or Unsatisfactory Report" form called a FUR. The pencil-checked FURs, originated at hundreds of operating activities, are mailed to a central point, converted to punch cards and then mailed directly to the prime contractor. When the originator also amplifies his pencil checks with worded comment or photographs then a 35mm transparent film reproduction of the entire report is included right in the punch card. Echo information of any resultant manufacturer's action goes back later to the originating activity to encourage future reports and to forestall that frustrating feeling that "nobody cares". The volume of FURs has already reached high levels and the early characteristic discouragement period has occurred but now seems to be receding.

The airframe manufacturer noted above has reported that the FUR system cut his average data feedback time to 40% of its former length. The volume of reports on comparable numbers of his aircraft increased by hundreds of times and the resulting comparative data helped him in the solution of basic problems. The FUR data enabled better comparative evaluation of the service capability of vendor items. He was aided by the fact that he already had in operation a system of tabulation and statistical analysis of related mass data at the time the rush of new FUR cards on Naval aircraft arrived.

In summary, it now seems more practicable to obtain mass information on the experience of the military consumer of aircraft. The information can be transmitted, sorted and initially analyzed by rapid modern mass methods. Selective analysis, leading to corrections or improvements, then becomes feasible. The net result of initial ventures looks profitable, at least in the aircraft field and after the initial expansion period is survived.